

SURVEILLANCE OF VEHICLE OBJECTS WITH AERIAL IMAGES USING LOCALIZATION AND POSTURE BASED LOCAL GRADIENT MODEL

¹R.C.KARPAGALAKSHMI, ² DR.D.TENSING

¹Assistant professor, Department of CSE, Tagore Institute of Engineering & Technology,
Salem, 636 112, Tamil Nadu, India.

²Director, School of Civil Engineering, Karunya University,
Coimbatore – 641 114, Tamil Nadu, India.

E-mail: ¹karpagalakshmir@gmail.com, ²dtensing@rediffmail.com

ABSTRACT

Traffic scenes in road network needs to be monitored in different dimensional perspective to address the issues generated at various time instances. One of the major challenges observed in traffic scenes in road network is to address the vehicle tracking at different dimensional perspective for various time instances. One such application that has received greater significance is the mechanism to develop a full proof vehicle traffic control scheme. Videos fitted on different proposition of the signal junction able to view at respective image positioning and the overall traffic may rise to abnormality. Many studies have been examined for vehicle traffic control scheme. The recently used scheme is model based on simple object recognition and localization of road vehicles based on the position and orientation of vehicle image data. But the drawback of the approach is that if the shape of the vehicle and its pose varies in multiple junction coordination, the model based recognition is an inefficient one. To overcome the issues, in this work we are going to implement surveillance image object recognition and localization using improved local gradient model. The vehicle-object shape recognition and pose recovery in the traffic junction is carried out for varied traffic densities. An experimental evaluation is carried out to estimate the performance of the proposed Surveillance of Vehicle Object Recognition and Localization (SVORL) using improved gradient model in terms of vehicle density, traffic junction points, and computation time and compared with an existing model based on simple object recognition and localization.

Keywords: *Vehicle Object Recognition, Object Localization, Improved Gradient Model, Ray Traced Templates, Road Extraction*

1. INTRODUCTION

As the residents in cities persistently rise, traffic in cities develops into crammed a lot more than city and communications preparation has estimated. To resolve this problem, supervising vehicles is one of the significant tasks. Generally, the field-based utensils like cameras mounted at permanent locations or weigh-in action sensors on the concrete are employed to examine traffic. Thus, vehicle recognition during remote sensing can be widely utilized to deal with assess fuel demand, traffic, guess traffic secretion and also significant for transport communications planning. The representation of road network is shown in figure 1.

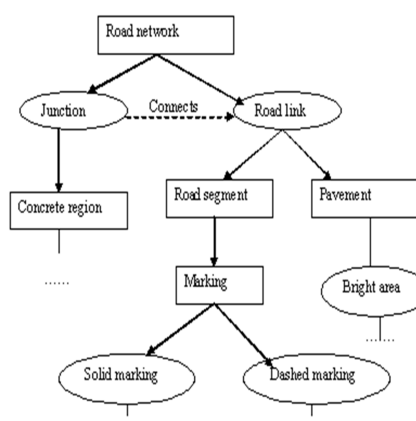


Figure 1: Road Network Model

The emergence of roads in digital metaphors supported on the ethereal compassion as well as on the declaration of then antenna. Particularly, a road can materialize in diverse ways depending on the level of surveillance. In images with common declaration (more than 2 m per pixel), roads emerge largely as lines, which institute a more or less opaque network. In disparity, in high-resolution images (less than 0.50 m per pixel) roads are portrayed as bright, more or less standardized lengthened areas with roughly invariable width and enclosed curvature. In this work, we are going to obtain the aerial imagery clear view through the exact position, location, angular view of the camera and the improved gradient model are implemented to estimate the recognition and localization of vehicle objects in the traffic scenes.

The contributions of Surveillance of Vehicle Object Recognition and Localization includes the following

1. To capture the correct posture of the image using ray traced templates
2. To recognize the vehicle object using improvised gradient information model recognize the object in motion
3. Using continual localization approach to determine the location and improve the vehicle tracking.

2. LITERATURE REVIEW

The vehicle object recognition and localization are an imperative survey in a recent study to estimate the moving vehicle objects in the road traffic junction boards. Many techniques address the problem of model-based object recognition. The study intends is to confine and distinguish road vehicles from monocular images or videos in standardized traffic scenes. A 3-D deformable vehicle model with 12 figure restrictions is place up as previous information, and its pose is resolute by three parameters, which are its location on the earth plane and its direction about the perpendicular axis underground-plane restraints [1].

A deformable motor vehicle representation seems to be a recovered choice to conquer the disadvantages declared. On one hand, it can exercise the rewards of model-based methods [12]. On the other hand, it can adjust itself to compact with most vehicles in certainty with vehicles of diverse types and poses efficiently [7]. A current development has been planned in [3] to realize vehicle arrangement based on a 2-D shape model,

which can contract efficiently with diverse noise, views, and occlusion.

As we can see, evolutionary calculating is a stochastic searching-based method, which has previously had superior requests in human body-motion analysis [4]. From one monocular image of standardized scenes comprising vehicles, we can improve both pose and shape parameters to appreciate the localization and appreciation of vehicles [10].

In [13], motivated by the object recognition the author presented a new coarse-to-fine global localization approach. The combined work of coarse and fine stages makes the system more fast and reliable which at the same time recovers the position of the camera by matrix decomposition. A multi-view approach was presented in [14] to solve the problem of occlusion using 2D constructs. A planar homographic occupancy constraint was developed to fuse foreground likelihood information from multiple views, to resolve occlusions and localize people on a reference scene plane. The work in [15] described a composition system to automatically learn the structured, hierarchical representation of object in an unsupervised way without the need of segmentation to be performed manually.

For affecting vehicles, we can haul out regions of attention by motion discovery and attain vehicle regions by action and shape information, which is fine depicted in [5]. For stationary vehicles, many geometric models can be employed to discover vehicles, which have been well recapitulated in [6]. The particulars of recognition are not the importance of the spring boxes of motor vehicle regions have previously been attained. The auto calibration method [9] is specified detail in [8].

3. PROPOSED ESTIMATING VEHICLE OBJECT RECOGNITION AND LOCALIZATION USING IMPROVED LOCAL GRADIENT MODEL

The proposed work is efficiently designed for identifying the vehicle object recognition and localization using an improved gradient model. The vehicle object and recognition is done based on aerial view of images. The aerial vehicle images are obtained by fixing a camera on junction board to get a clear extraction of road view to identify the vehicle densities and postures. The proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model [EVORL] is worked based on following two

phases. The architecture diagram of the proposed SVORL is shown in figure 2.

The first process is to positioning the camera based on the junction board, so that we can obtain the exact aerial view of road in a clear manner. To approximate surveillance of individual vehicle aerial imagery in elevated resolution aerial photographs, use rounded templates consequent from an optical tree circlet model and the geometry at image attainment. A direct analysis method comprises of identifying local maxima of the association function for the consecutively transposed templates over the whole image.

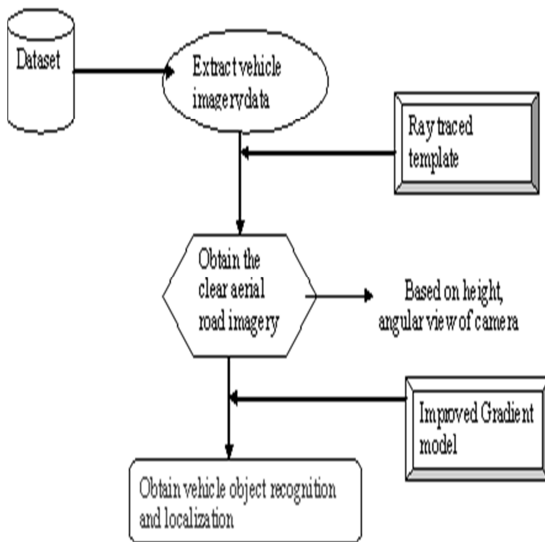


Figure 2: Architecture Diagram Of The Proposed SVORL

The second phase describes the process of identifying the protuberance of the vehicle model and image data, which is integrated into a new evolutionary computing framework to guess the surveillance image shape constraints and pose constraints by iterative evolution.

From the figure 3.1, it is being observed that the vehicle object and recognition are efficiently done based on the position and localization of the camera fixed on junction board. The position and localization of camera is done through ray traced templates to obtain clear aerial road imagery and should observe the road till the end. After fixing the camera based on its angular view, the vehicle object recognition and localization are done using the improved gradient model.

3.1 Fixing Camera Using Ray Tracer Method for Surveillance

The ray tracer method is used here on aerial road imagery templates modified to the geometry at image attainment. It should be mainly important for images attained from angles where the vehicle objects are back-lighted by the reflection, a position that potentially may provide the most precise vehicle object summit position approximates for elegant and other objects with analogous morphology.

The camera is identified to be a pinhole camera. Its position, decision, the direction of the visual axis and the “up” bearing in the image plane are calculated from the image refinement parameters. The crucial camera geometry is defined in such a way as, A summit in image coordinates $(a, b)^T$ can be distorted to a ray (or a 3D vector), $(m, n, 1)^T$, imminent from the camera origin $(0, 0, 0)^T$ and transient through the image plane $O = 1$ (See Figure 3.2). The ray $(m, n, 1)^T$ can be changed (revolved and rescaled) into world coordinates, $(r_M, r_N, r_O)^T$, where

$$\begin{Bmatrix} r_M \\ r_N \\ r_O \end{Bmatrix} = R^{-1} \begin{Bmatrix} m \\ n \\ 1 \end{Bmatrix} \quad (1)$$

where R is the revolution matrix (world to camera). For expediency, we imagine in the following that the camera and vehicle harmonizes (positions) are similar in the world coordinate frame as adding up another conversion is straightforward. When the vehicle’s 3D location in world coordinates is $(T_M, T_N, h)^T$ the object coordinates are

$$\begin{pmatrix} M \\ N \\ O \end{pmatrix} = \begin{pmatrix} T_M \\ T_N \\ h \end{pmatrix} - \frac{h}{rz} \begin{pmatrix} r_M \\ r_N \\ r_O \end{pmatrix} \quad (2)$$

Assuming that the target is on flat ground ($Z = 0$).

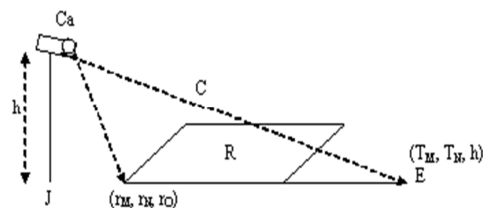


Figure 3: Fixing camera position

From the figure 3, the position of the camera is fixed based on the height and angular view of the images. The camera Ca is positioned based on the height H of the junction board J. The camera view should focus the aerial vehicle imagery till the end i.e., it should cover the road at end to view the vehicle traffic in a clear manner using the world real co-ordinates. If C is the path of the camera ray through the vehicle object, then the view reflected by the camera ray is termed as

$$V(ca) = \sum_{i \in \text{length of } E} R(i, C) K_{\text{refl}} \quad (3)$$

where V(ca) specifies the view of camera from junction board

E- End of the road

C- Path of camera ray

K refl – constant obtaining the total reflection aerial imagery view

The ray-tracer is employed to construct a spherical pattern image of a vehicle object, as it would appear along with the model if it was positioned in the centre of the junction board. For large stands it might be essential to construct numerous templates for diverse locations inside the stand (preferably a diverse template would be formed for each pixel). The template is centered on the top of the (synthetic) camera and has a radius corresponding to how much of each vehicle object is expected to be visible in the image (this of course based on screening crown geometry, angle, and tree density inside the stand). For each image pixel within the stand the association between the template and a consequent rounded window centered on the pixel is evaluated. The constructive local maxima of the resultant association image communicate to the locations that (nearby) be similar to the template most directly and therefore are candidates for camera top positions.

3.2 Vehicle Object Recognition and Localization Using Improved Gradient Model

Normally, vehicles are synthetic objects of previous shape information. Projections of a vehicle in the image flat surface are affluent in lines of the three orientations $(\vec{O}_1, \vec{O}_2, \vec{O}_3)$ which communicate to three orthogonal directions. We can estimate these three orientations in the image plane based on image gradients and use angles among them as constraints to estimate the initialization value of θ . Since the axes are concurrent to be vertical to the ground flat surface, the projection of can be estimated in the image

plane based on the calibrated camera. There are only two peaks in the histogram of oriented gradients (HOG). By comparing these two peaks to the calculated, we can confirm that they are corresponding to the respective orientations $(\vec{O}_1, \vec{O}_2, \vec{O}_3)$. With a second-step HOG in the locality of every peak, we can attain added more precise inference of angles among and can be considered merely, and the initialization of is achieved by [1],

$$\cos \gamma = \frac{\vec{O}_1 \cdot \vec{O}_2}{|\vec{O}_1| |\vec{O}_2|} \quad (4)$$

With the pose constraints initialized, the vehicle model can be anticipated into the image flat surface to contest with image data. A precise and proficient method is mandatory for fitness valuation among the projection of the model and image data. With shape and pose constraints initialized, we can develop the wire-frame vehicle representation onto the image level surface to figure a sequence of evident projected line segments. For each evident predictable line segment whose direction is dispensed as with length of in the image level surface, we form a virtual rectangle, as shown in Figure 3. If the line fits the image data well, the gradient directions of pixels with large gradient magnitude values in the rectangle should specify on the vertical course of the projected line. Consequently, we can approximate the fitness score from the inclined information of all pixels inside the rectangle as follow,

$$F = m(a, b) \cdot \sin(\beta(a, b) - \alpha) \quad (5)$$

Where F - Fitness score

m(a, b) - gradient magnitude

$\beta(a, b)$ – Orientation of pixel of vehicle image data

Like this, the fitness score of the anticipated line segment is considered. Compared with an existing MOR, we use inclined information from each pixel in the local regions in place of estimating directional imitative from modeled image points. This strategy provides effectiveness but adds strength to noise and occlusion. To recover the shape and pose of the given image, the proposed SVORL framework first set the pose parameters and the value range which is used to obtain the correct pose of the vehicle object. With the value range of shape and pose parameters, the individual vehicle object parameter space is obtained and recovered and identified the path of the vehicle which is yet to be mover from the road traffic junction. The algorithm below describes the

process of identifying the vehicle object recognition and localization using improved gradient model.

4. EXPERIMENTAL EVALUATION

The proposed estimating vehicle object recognition and localization using improved local gradient model is efficiently designed and implemented in Java to identify the surveillance aerial images in a correct point of view and identify the exact location and position of the vehicle objection standing in the junction board. At first set up, the surveillance images are captured based on the position, location, angular view of the images using ray traced scheme. Then the vehicle object recognition and the localization are identified using the improved gradient model which exactly identifies the vehicles drag on the traffic phase. The performance of the proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model [SVORL] is measured in terms of vehicle density.

5. RESULTS AND DISCUSSION

In this work, we have seen how the vehicle image object recognition and localization is done with improved local gradient model. The vehicle-object shape recognition and pose recovery in the traffic junction is carried out for varied traffic densities. Compared to an existing model based on simple object recognition [MOR] and localization, the proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model [EVORL] obtained an efficient outcome by simply providing a vehicle object and recognition using improved gradient model. The below table and graph describes the performance of the proposed estimating vehicle object recognition and localization using improved local gradient model.

Table 1: Vehicle Density Vs. Object Recognition And Localization

Vehicle density	Object recognition and localization (%)	
	Proposed SVORL	Existing MOR
10	54	25
20	66	32
30	78	40
40	84	54
50	95	60

The table (table 1) describes the vehicle object recognition and localization based on

obtaining clear surveillance images. The outcome of the proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model [SVORL] is compared with existing model based on simple object recognition [MOR].

Figure 4 describes the vehicle object recognition and localization based on the vehicle densities in the road traffic signal. In practice, we can gather a set of distinctive vehicles for every type and compute their standard shape analogous to the model individual of that kind in the shape parameter space. For vehicle recognition, the proposed SVORL improve primary the shape parameters of the vehicle, which analogous to an entity in the shape restriction space, and contrast its distance to form individuals of all types. The vehicle is renowned as the type whose representation individual is contiguous to the improved shape parameters. The set comprises vehicles of diverse types, diverse view angles, in the case of trees, bicycles, pedestrians, and other vehicles. The scenes are regulated virtually supported on the motion and emergence information of affecting objects in videos with camera elevation considered as the only one user input.

Compared to an existing MOR, the proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model obtains a clear vehicle object recognition and localization.

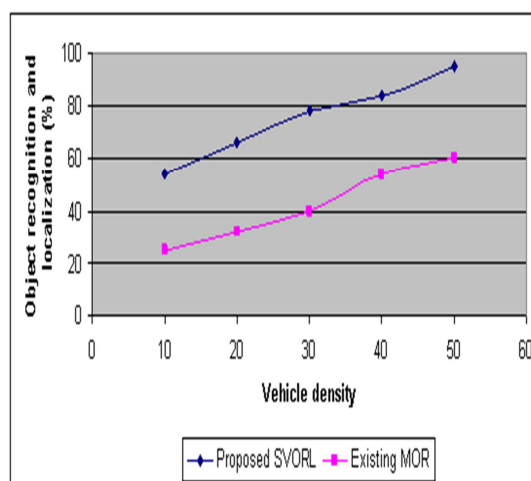


Figure 4: Vehicle Density Vs. Object Recognition And Localization

6. CONCLUSION

A novel approach has been described here for the model supported localization and recognition of vehicles from monocular images by fitting the camera in an exact position to obtain a surveillance images. A deformable model camera is set up using ray traced method, and an efficient method based on the image gradient is proposed to assess fitness among the protrusion of the vehicle model and image data. An evolutionary structure is approved to produce a huge number of models based on the proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model. The proposed Surveillance of Vehicle Object Recognition and Localization using improved local gradient model cannot only comprehend vehicle localization and recognition but also recover the real shape of diverse kinds of vehicles. Experimental results reveal the effectiveness of the proposed SVORL with diverse vehicles, diverse poses. To further improve the work to extract the aerial view of road images obtained from any digital cameras, a strong geometric model can be designed which remains an open issue and can be explored as future area.

REFERENCES:

- [1] Zhaoxiang Zhang, Tieniu Tan et. Al., 'Three-Dimensional Deformable-Model-Based Localization and Recognition of Road Vehicles', *IEEE TRANSACTIONS ON IMAGE PROCESSING*, VOL. 21, NO. 1, JANUARY 2012.
- [2] Bernd-Michael Wolf (né Straub) · Christian Heipke, "Automatic extraction and delineation of single trees from remote sensing data", *Machine Vision and Applications (2007)* 18:317–330 DOI 10.1007/s00138-006-0064-9.
- [3] Zu Whan Kim, and Raja Sengupta, "Target Detection and Position Likelihood using an aerial image sensor", *JCR/IEEE (2008)*, p. 59-64.
- [4] Y. Li, L. Gu, and T. Kanade, "A robust shape model for multi-view car alignment," in *Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recognition*, 2009, pp. 2466–2473.
- [5] J. Deutscher and I. Reid, "Articulated body motion capture by stochastic search," *Int. J. Comput. Vis.*, vol. 61, no. 2, pp. 185–205, Feb. 2005.
- [6] Z. X. Zhang, Y. Cai, K. Huang, and T. Tan, "Real-time moving object classification with automatic scene division," in *Proc. IEEE Conf. Image Process.*, 2007, pp. V-149–V-152.
- [7] Z. Zhang, K. Huang, and T. Tan, "3D model based vehicle localization by optimizing local gradient based fitness evaluation," in *Proc. Int. Conf. Pattern Recog.*, 2008, pp. 1–4.
- [8] Z. Zhang, K. Huang, T. Tan, and Y. Wang, "3D model based vehicle tracking using gradient based fitness evaluation under particle filter framework," in *Proc. Int. Conf. Pattern Recog.*, 2010, pp. 1771–1774.
- [9] Z. Zhang, M. Li, K. Huang, and T. Tan, "Practical camera auto-calibration based on object appearance and motion for traffic scene surveillance," in *Proc. Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2008, pp. 1–8.
- [10] Farah, I.R. et. Al., "Multiapproach System Based on Fusion of Multispectral Images for Land-Cover Classification", *IEEE Transactions on Geoscience and Remote Sensing*, 2008.
- [11] Zongli Jiang et. Al., "Applying SAR interferometric coherence to outline debris-covered glacier", *19th International Conference on Geoinformatics*, 2011.
- [12] Xiao Xiang Zhu et. Al., "A fundamental bound for super-resolution with application to 3D SAR imaging", *Urban Remote Sensing Event (JURSE)*, 2011 Joint.
- [13] Junqiu Wang, Hongbin Zha, Cipolla, R., "Coarse-to-fine vision-based localization by indexing scale-Invariant features", *Systems, Man, and Cybernetics, Part B: IEEE Transactions on Cybernetics*, Volume : 36, Issue : 2, Apr 2006.
- [14] Khan, S.M., Shah, M., "Tracking Multiple Occluding People by Localizing on Multiple Scene Planes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume : 31, Issue: 3, Mar 2009.
- [15] Ommer, B., Buhmann, J.M., "Learning the Compositional Nature of Visual Object Categories for Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume:32, Issue: 3, Mar 2010.