

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Vision-Based Road Detection by Monte Carlo Method

¹W. Yanqing, ¹C. Deyun, ²S. Chaoxia and ¹W. Peidong

¹School of Computer Science and Technology,
Harbin University of Science and Technology, Harbin,
Heilongjiang Province, 150080, China

²School of Computer Science and Technology,
Nanjing University of Science and Technology, Nanjing,
Jiangsu Province, 210094, China

Abstract: A novel vision-based road detection method is proposed in this study to realize visual guiding navigation for Autonomous Land Vehicles (ALV). The captured image was first segmented into road region and non-road region by using Otsu thresholding algorithm. Subsequently, the extracted Canny edges located in deep position of road region and non-road region would be removed by filtering so that the recognition of road boundary could not be affected by mussy edges existed in the image. In order to improve the performance of road detection, the dynamics of ALV and the Monte Carlo Method was taken into account to associate the possible road boundary in different image frames. The method proposed in this study is robust against strong shadows, surface dilapidation and illumination variations. The real world experiment on road detection has demonstrated that the proposed method is feasible and valid.

Key words: Road detection, Otsu thresholding method, edge extraction, Monte Carlo method, image segmentation, the intelligent transport systems

INTRODUCTION

Road detection techniques have been extensively developed due to their widely applications in driver assistance systems, visual navigation of Autonomous Land Vehicles (ALV), the Intelligent Transport Systems (ITS) and other corresponding researching fields (He *et al.*, 2004).

Due to the fact that camera can acquire more abundant information from its surroundings and at the same time supply a cheaper way for autonomous navigation, several vision based road-following application systems have been developed. Some researchers focus their works on structured roads, in which clear marks can be recognized. The Generic Obstacle and Lane-Detection (GOLD) system (Bertozzi and Broggi, 1998) at Parma University, Italy, reprojects the image ahead of the vehicle onto the ground plane and extracts left-middle-right groups of road markings. The Yet Another Road Follower (YARF) system (Kluge and Thorpe, 1992) pays even more attention to the modeling of road geometry and thereby estimates road-model parameters based on the road-marking extraction result.

Obviously, these methods are not appropriate in many real-life situations in which the roads have no clear road markings, such as roads in urban, campus and country environment. Therefore, some researchers propose the road detection methods those are applicable to unstructured roads. The Supervised Classification Applied to Road Following (SCARF) (Crisman and Thorpe, 1993) and unsupervised clustering applied to road following (UNSCARF) systems (Crisman and Thorpe, 1991) at Carnegie Mellon University, have been built to deal with road in parks, with no lane markings and considerable difference in color between the road area and surroundings.

Four typical methods were proposed to realize robust road detection for the unstructured road which have no lane markings, may have degraded surfaces and edges and may be partially obscured by strong shadows:

Methods based on road features: Such methods, i.e., feature clustering, threshold segmentation or region growing, usually catch the key point on color, intensity or texture information between the road region and the

non-road region, are usually insensitive to road shapes, but sensitive to illumination, shadows and water marks.

Methods based on road models: Such methods often assume a road model first, then, find the fittest model according to original images. Although, these methods can detect the road more integrally, it may be invalid when facing diverse road shapes.

Methods based on Artificial Neural Networks (ANN) (Conrad and Foedisch, 2003): Such methods are independent of road markings and lanes. During a short initialization step, feature data is automatically collected based on the conforming road structure in images captured from the driver's point of view. Then, the new Neural Network is trained and applied to the new environment. This procedure allows the system to detect road adaptively and supplies a promising way for road recognition and environmental learning. However, these methods often fail when current road is dissimilar to that of the training set.

Methods based on vanishing points (Se and Brady, 2000; Simond and Rives, 2003): Generally, the key to the approach is to use a voting procedure like a Hough transform on edge-detected line segments to find the intersection point of different road edges (Nguyen *et al.*, 2009). Similar grouping strategies have also been investigated outside the context of roads, such as in urban and indoor environment rich in straight lines, in conjunction with a more general analysis of repeated elements and patterns viewed under perspective (Turina *et al.*, 2001; Schaffalitzky and Zisserman, 2000).

When performing vision-based road detection, the input is a camera image. It can be a grey image or a color image. To locate the road in the image some assumptions regarding the image information are made:

Firstly, the road region differs from the non-road region in color or intensity features.

Secondly, most part of the road has similar features and certain different feature caused by watermark, strong shadow and surface dilapidation is also permitted.

Based on the assumptions mentioned above, the road region was first recognized against its surrounds by using Otsu thresholding algorithm in part II. In part III, the Canny edges extracted in grey images would be filtered in the road region and non-road region. The Monte Carlo method and the dynamics of ALV would be discussed in part IV. After the road detection experiments in different scenes and different illuminations, some conclusions would be given in part V.

ROAD REGION RECOGNITION BASED ON OTSU THRESHOLDING METHOD

Region recognition can be handled by popular thresholding algorithm such as Maximum Entropy, Invariant Moment and Otsu thresholding method. Because Otsu thresholding method supplies a more satisfactory performance in image segmentation, it was used to overcome the negative impacts caused by environmental variation (Otsu, 1979).

In order to utilize the most important color information of the color image, the candidate color channel that was dominant in certain color space was selected to realize image segmentation.

$$C_c = \underset{(R, G, B)}{\operatorname{argmax}}(N_R, N_G, N_B) \quad (1)$$

Equation 1 indicates the principle of the color channel selection, in which CC means the color channel; NR means the number of the dominant color channel in certain referenced regions. Because the camera is installed in a fixed place, it is a reasonable hypothesis that certain region can be regarded as road region. The dominant color channel of the region will be selected adaptively to calculate the histogram of the image. Otsu thresholding method was then used to segment the whole image into road region and non-road region adaptively.

ROAD BOUNDARY DETECTION BASED ON CANNY EDGE

Though Otsu thresholding method can relieve such negative impacts as illumination variations, strong shadows or surface dilapidation to some extent, it often mistakes certain part of non-road region for road region due to the similar feature of nearby buildings. Therefore, the Canny edge will be extracted and regarded as the constraint of the road region so that both the accurate road region and the road boundary can be acquired simultaneously.

During the course of Canny edge detection, we have to handle the dilemma as follows: The less the parameter of edge strength selected, the more possible road boundary will be extracted, but it is difficult to extract the due road boundary from so many Canny edges; Increasing the parameter of edge strength will reduce the quantity of the Canny edges naturally, but some weak edges that located on road boundaries may be neglected inevitably.

In order to resolve the dilemma mentioned above, we combine the road region and non-road region with the

Canny edges. Firstly, appropriate edge strength will be selected to produce enough strong and weak Canny edges. Secondly, redundant edges lie in deep non-road region or in deep road region will be removed.

$$E(x, y) = \begin{cases} \overline{E(x, y) \prod_{t=-r}^r R(x+t, y+t) \prod_{s=-r}^r R(x+s, y-s)} & R(x, y) = 0 \\ E(x, y) \prod_{t=-r}^r R(x+t, y+t) \prod_{s=-r}^r R(x+s, y-s) & R(x, y) = 1 \end{cases} \quad (2)$$

The principle of Canny edge filtering is described in Eq. 2, where, $E(x, y)$ means whether the pixel (x, y) is the edge or not (1: edge; 0: non-edge); $R(x, y)$ means whether the pixel (x, y) is the road region or not (1: road region; 0: non-road region); r is the depth threshold and the upper line in the right part of the Eq. 2 means the negative operator.

After the redundant Canny edges have been filtered by road and non-road region, the remaining edges will be used to produce several lines by Hough Transform. One or two of those lines will be selected to fit nearby left or right road boundary by considering the distance from the vehicle to the line, the mean strength of the Hough line, as well as each line's orientations.

MONTE CARLO METHOD

For most small curvature roads, nearby road boundary can be fitted by straight line. Therefore, the Hough transform was used to detect the road boundary in this work for the reason that it is invariant to pixel position in an image. In Hough space, each edge segment is responding to a certain line and each line can be represent with $L(\rho, \theta)$, where, ρ is the distance from the line to ALV and θ is the angle that the line makes to the Ox-coordinate axis. Taking into account the dynamics model of the vehicle, ALV can predict the variation of $L(\rho, \theta)$ with its moving.

Dynamics of ALV: The dynamics of ALV is expressed in Eq. 3, where, $(\Delta x_k, \Delta y_k, \Delta \theta_k)$ is the variation of the vehicle's pose (position and orientation) at time k , (v_k, ω_k) is the velocity space composed of linear velocity and angular velocity ω_k .

$$\begin{bmatrix} \Delta x_k \\ \Delta y_k \\ \Delta \theta_k \end{bmatrix} = \begin{bmatrix} \cos(\omega_k t_k) & 0 \\ \sin(\omega_k t_k) & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} v_k t_k \\ \omega_k t_k \end{bmatrix} \quad (3)$$

Considering most vehicles have the capability of acquiring its velocity information rather than its position

$$\begin{cases} \rho_k = \rho_{k-1} + v_{k-1} \cdot t_{k-1} \cos(\pi - \theta_{k-1} + \Delta \theta_{k-1}) \\ \theta_k = \theta_{k-1} + \Delta \theta_{k-1} \end{cases} \quad (4)$$

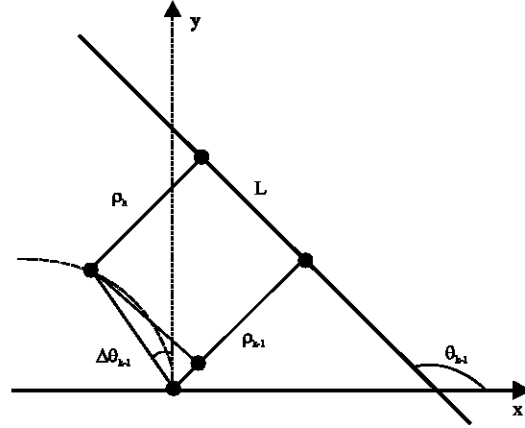


Fig. 1: Prediction of $L(\rho, \theta)$

and orientation in certain global reference frame, only the distance ρ and the angle θ of a possible road boundary in the vehicle's reference frame are calculated in the dynamics of ALV.

When ALV moves forward with velocity (v_k, ω_k) at time step k , based on Fig. 1 and Eq. 3, the prediction of $L(\rho, \theta)$ at time step k can be calculated by Eq. 4.

Monte Carlo method: Provided the belief of the system state at time t , L_t , is represented by $Bel(L_t)$, apparently, $Bel(L_t)$ also means the possibility that the i th line l_i in Hough space belongs to left or right road boundary. According to the Bayes law and Markov assumption (past and future data are independent if the current state of the system is known), $Bel(L_t)$ can be updated by both the perception model and the motion model:

$$Bel(L_t) = \eta p(o_t | L_t) \int p(L_t | L_{t-1}, u_{t-1}) Bel(L_{t-1}) dL_{t-1} \quad (5)$$

where, η is a normalizing constant. Perceptual model $p(o_t | L_t)$ means the valuation of L_t purely depending on the perception information captured by vision system on board. In motion model $p(L_t | L_{t-1}, u_{t-1})$, the estimation of current state L_t depends not only on the belief of previous state L_{t-1} , but also on the past neighboring executed action u_{t-1} , that could be calculated in Eq. 3.

To acquire $L(\rho, \theta)$ in real-world coordinates, the corresponding relation between the Hough line in image and the road boundary in actual surroundings should be considered. However, limited to the content of the study, this part will not be discussed.

Representing the continuous probability density function (pdf) by a set of n weighted samples is computationally more efficient and proves satisfactory also for representation of multi-modal and non-Gaussian belief states if the number n is sufficiently large:

$$\text{Bel}(L_t) = \{L_t^{(i)}, \varpi_t^{(i)}\}_{i=1,2,\dots,n} \quad (6)$$

where, $L_t^{(i)}$ is a sample state of the random variable L_t and the non-negative numerical factor $\varpi_t^{(i)}$ is an importance weight factor representing probability measure of each particle. The $L_t^{(i)}$ is defined as $\{\rho_t^{(i)}, \theta_t^{(i)}\}$ with the probability $\varpi_t^{(i)}$ associated to this hypothesis.

Traditional Monte Carlo method at each time step comprises of three steps:

- **Resampling:** Resample N samples randomly from L_{t-1} according to the distribution defined by ϖ_{t-1} . A group of particles that received best probability estimates from the previous step have the high possibilities to be resampled from the whole particle set
- **Importance sampling:** Sample state $L_t^{(i)}$ from $p(L_t | L_{t-1}, u_{t-1})$, for each of the N possible state $L_t^{(i)}$ and evaluate the importance factor $\varpi_t^{(i)} = p(o_t | L_t^{(i)})$
- **Summary:** Normalize the important factors $\varpi_t^{(i)} = \varpi_t^{(i)} / \sum_{i=1}^N \varpi_t^{(i)}$ and calculate the statistic property of sample set L_t to estimate the distance and the orientation of the possible road boundary

Since, samples are actually drawn from a proposal density, if the observation density moves into one of the tails of the proposal density, most of the samples' non-normalized importance factors will be small. In this case, a large sample size is needed to represent the true posterior density to ensure stable and precise tracking of the road boundary. Another problem is that samples often too quickly converge to a single, high likelihood pose. This might be undesirable in the environments have many line segment similar to the road boundary. To make the samples represent the posterior density better, proposed mixture Monte Carlo method, but it needs much additional computation in the sampling process (Thrun *et al.*, 2001). To improve the efficiency of Monte Carlo method, methods adjusting sample size adaptively over time are proposed, but they increase the probability of premature convergence (Fox, 2003).

In this study, a new version of Monte Carlo method is proposed to overcome those limitations above. Different operation on particle sample and evaluation has been executed. At the step of importance sampling, the sample state $L_t^{(i)}$ will be produced from the perception model $p(o_t | L_t^{(i)})$ and the selected N possible state $L_{t-1}^{(i)}$ will be evaluated by the importance factor $\varpi_t^{(i)} = p(L_t^{(i)} | L_{t-1}^{(i)}, u_{t-1})$ in motion model.

For lack of the priori knowledge of the road boundary such as the quantity of the lanes, the distance and the orientation etc., we approximate the perceptual model as Eq. 7, where, $R(x_{jk}^{(i)}, y_{jk}^{(i)})$ is described in Eq. 2:

$$p(o_t | L_t^{(i)}) = \frac{\sum_{j=0}^m \left| \sum_{k=0}^r R(x_{jk}^{(i)}, y_{jk}^{(i)}) - \sum_{k=-r}^0 R(x_{jk}^{(i)}, y_{jk}^{(i)}) \right|}{\sum_{i=0}^n \left(\sum_{j=0}^m \left| \sum_{k=0}^r R(x_{jk}^{(i)}, y_{jk}^{(i)}) - \sum_{k=-r}^0 R(x_{jk}^{(i)}, y_{jk}^{(i)}) \right| \right)} \quad (7)$$

The above likelihood function has the advantage of being quick to evaluate. This is important for real-time operation since, it needs to be computed for every particle at each time step. However, its discrete form means that care must be taken in setting the threshold. Setting it too high results in reduced discrimination ; Reducing it increases discrimination, but requires greater coverage of the sampling space so that enough samples are drawn around significant modes. Unfortunately, increasing the number of samples to achieve this quickly becomes prohibitive. To address this, we use particle annealing, in which samples are iteratively focused onto potential modes, whilst gradually reducing the threshold (Deutscher *et al.*, 2000).

At each time step, we use a large value for the threshold to obtain an initial set of weighted particles. Annealing then proceeds by resampling from this set using a uniform distribution with a smaller width than that used to propagate the samples from the previous time step. An updated set of weighted particles is then obtained using a smaller value of the threshold in the likelihood and this initializes the next annealing step. The process continues using smaller and smaller values of annealing velocity and the threshold, resulting in greater concentration of the samples around modes in the state space.

Due to the fact that the particle sampling takes place in perception model rather than in motion model, we can adjust the size of the particles adaptively to the quantity of the Hough line detected in perception model. Therefore, the proposed method is robust to the variation of the environment: when the surroundings have too much disturbance, the increasing of the particles precise the tracking of the due road boundary; On the contrary, if the background is becoming simple, the decreasing of the particle means the quick reaction to the environment.

RESULTS

To validate the proposed algorithm in this study, several experiments on road detection have been executed

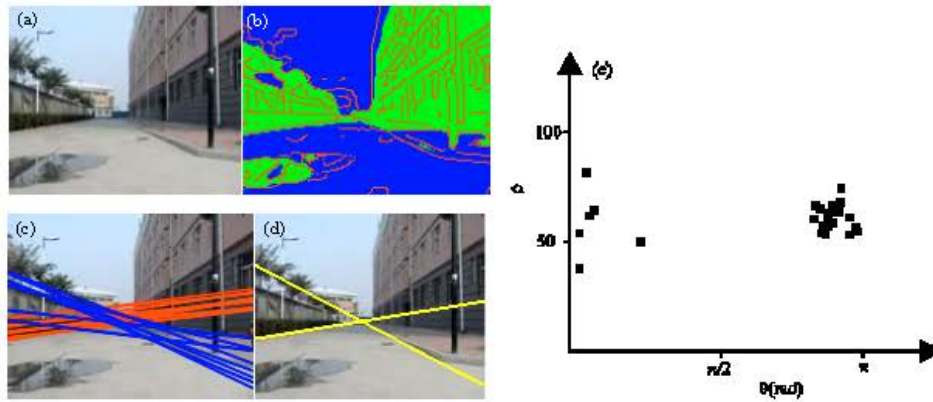


Fig. 2: Negative effects caused by watermark. (a) Original image, (b) Otsu-edge image, (c) candidate Hough line, (d) result of road boundary and (e) particle distribution

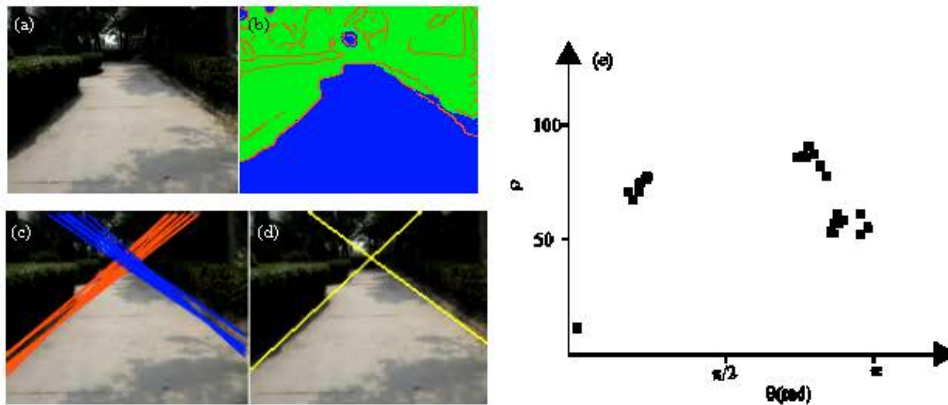


Fig. 3: Road detection under the condition of shadows, surface dilapidation. (a) Original image, (b) OTSU-edge image, (c) candidate Hough line, (d) result of road boundary and (e) particle distribution

in campus roads. The images were acquired in different scenes and different illuminations so that the proposed method can be assessed appropriately.

Experiments on road detection: The Negative impacts such as the illumination variation, the strong shadows, the surface dilapidation as well as the nearby buildings that have similar intensity features have tested by the single camera installed on the ALV.

Figure 2a-e showed us the negative effects result from watermark. The adaptive selection of the particle size is visually shown in Fig. 2, in which the particle distribution shown in Fig. 2e is corresponding to the Hough lines shown in Fig. 2c. In Fig. 2e, θ means the orientation between the Hough line and the positive x-coordinates and ρ means the image distance from the Hough line to the middle bottom point of the image.

Though the watermark in Fig. 2b was mistaken for non-road region by Otsu thresholding method (green region), the road boundary in Fig. 2d could be recognized among the candidate Hough lines shown in Fig. 2c by virtue of Monte Carlo method. The distribution of the particle was shown in Fig. 2e, where all the particles huddle in different regions.

Figure 3a-e showed us the experimental results when shadows, surface dilapidation happen simultaneously. Fig. 3b indicated that when most Canny edges on both background and the road surface were removed by Otsu thresholding image, the detail of the weak road boundary was well preserved.

In Fig. 4a-e both the road surface and the nearby buildings had the similar color and intensity features. By combining the road region segmentation, the Monte Carlo method with the road boundary extraction, the road's boundary could be recognized correctly.

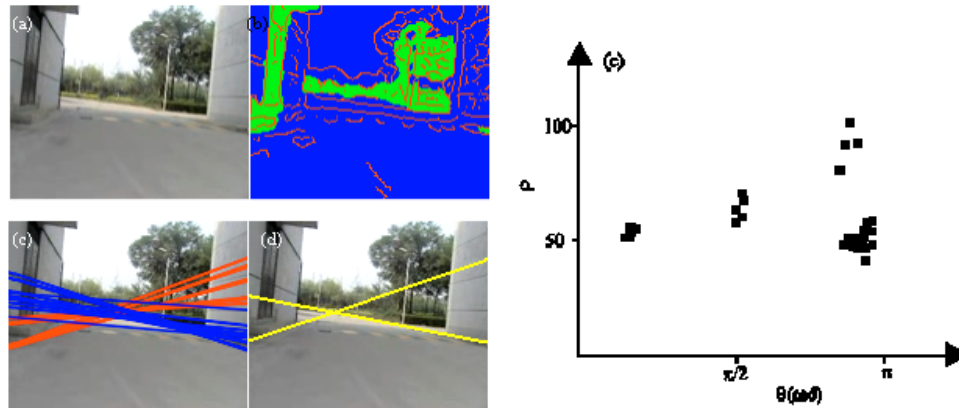


Fig. 4: Negative impacts result from similar scene. (a) Original image, (b) Otsu-edge image, (c) candidate Hough line, (d) result of road boundary and (e) particle distribution

CONCLUSION

A novel road detection algorithm that combined the road region segmentation, the Monte Carlo method with the road boundary extraction was proposed in this study. Road detection experiments illustrate that the proposed method is not only robust to such negative impacts as illumination variation, strong shadows, surface dilapidation and water marks, but also it can meet the requirement of real-world navigation for autonomous land vehicle.

ACKNOWLEDGMENTS

This study has been supported by Major Program of National Natural Science Foundation of China (grant No. 90820306) and supported by the Special funds for the Research on Scientific Creation and Talent People (grant No. 2008RFQXG067) and Supported by Educational Commission Science and Technology of Heilongjiang Province of China (grant No. 11541040; 11541050).

REFERENCES

Bertozzi, M and A. Broggi, 1998. GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Trans. Image Process.*, 7: 62-81.

Conrad, P. and M. Foedisch, 2003. Performance evaluation of color based road detection using neural nets and support vector machines. *Proceedings of the Applied Imagery Pattern Recognition Workshop*, Oct. 15-17, Washington, DC., pp: 157-160.

Crisman, J.D. and C.E. Thorpe, 1991. UNSCARF: a color vision system for the detection of unstructured Roads. *Proc. IEEE Int. Conf. Robotics Automation IEEE Comput. Soc.*, 3: 2496-2501.

Crisman, J.D. and C.E. Thorpe, 1993. SCARF: A color vision system that tracks roads and intersections. *IEEE Trans. Robotics Automation*, 9: 49-58.

Deutscher, J., A. Blake and I. Reid, 2000. Articulated body motion capture by annealed particle filtering. *Proc. Int. Conf. Comput. Vision Pattern Recognition*, 2: 126-133.

Fox, D., 2003. Adapting the sample size in particle filters through KLD-sampling. *Int. J. Robot. Res.*, 22: 985-1004.

He, Y., H. Wang and B. Zhang, 2004. Color-based road detection in urban traffic scenes. *IEEE Trans. Intel. Transport. Syst.*, 1: 309-318.

Kluge, K. and C. Thorpe, 1992. Representation and recovery of road geometry in YARF. *Proceedings of the IEEE Symposium Intelligent Vehicles*, June 29-July 1, Carnegie Mellon Univ., Pittsburgh, PA., pp: 114-119.

Nguyen, T.T., X.D. Pham and J.W. Jeon, 2009. Rectangular object tracking based on standard hough transform. *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, Feb. 22-25, IEEE., pp: 2098-2013.

Otsu, N., 1979. A threshold selection method from grey level histogram. *IEEE Trans. Syst. Man Cybern.*, 9: 62-66.

Schaffalitzky, F. and A. Zisserman, 2000. Planar grouping for automatic detection of vanishing lines and points. *Image Vision Comput.*, 9: 647-658.

- Se, S. and M. Brady, 2000. Vision-based detection of staircases. Proc. 4th Asian Conf. Comp. Vision ACCV, 1: 535-540.
- Simond, N. and P. Rives, 2003. Homography from a vanishing point in urban scenes. Proc. Int. Conf. Intel. Robots Syst., 1: 1005-1010.
- Thrun, S., D. Fox and W. Burgard, 2001. Robust monte carlo localization for mobile robots. Artificial Intel., 128: 99-141.
- Turina, A., T. Tuytelaars and L. van Gool, 2001. Efficient grouping under perspective skew. Proc. Comput. Vision Pattern Recognition, 1: I-247-I-254.