A yaw and tilt invariant vehicluar ego-position model

James D.B. Nelson, Shan Fu .

February 22, 2005

Abstract

An analytical relationship is derived between the angles of two parallel road markings, projected onto an image plane, and the lateral position of the host vehicle. The resulting model proves that the relative position of a vehicle can be determined by the two lane angles alone, independent of camera height, yaw, and tilt. Furthermore, we find an optimally stable set of values for the measurements of height, yaw, and tilt with respect to the relative lateral position metric. Based on the Hough transform, a lane detection method is constructed to quantitatively validate the model.

1 Introduction

Measuring the lateral position of a car with onboard computer vision is not new. Owing to the passive nature and rich information content, computer vision has become a common method for the measurement of lateral position. The two principal motivations of such work are the applications to automatic vehicle guidance [1, 7, 10, 14, 18] and lane departure warning systems [6, 8, 21, 22].

However, the lateral position of a car is also indicative of driver attributes such as workload, ability, awareness, and aggression. To this end, there is growing interest in the application of lateral position measures to the study of driver behaviour [5, 11, 17]. Although our work was carried out in the context of driver behaviour studies, the application to automatic guidance and departure warning is naturally inferred.

Regardless of application, since a road model and hardware configuration are common prerequisites of state-of-the-art lane position systems, some kind of calibration procedure or a priori information about the hardware is often needed [1, 2, 10, 15, 16]. Furthermore, internal camera parameters, such as optical axis, are sometimes taken for granted.

On the other hand, we introduce a solution that obviates many such calibration requirements. In particular, unlike previous work, our method is invariant to camera height, yaw, and tilt.

The main thrust of our work is to find a relationship between the lateral position of the car and the angles that the two lane markings make in the image plane. Although based upon elementary geometry, this idea has, surprisingly, been overlooked by previous research. Most notable is the extensive model by Dickmanns and Mysliwetz [4], often cited in the literature.

In 2002, although Woong Lee [21] noted that lane angles and lane position were related, no attempt was made to find an analytical expression. A year later, Woong Lee et al. [22] came even closer and suggested that if the ratio of the two lane angles was too large or too small, then lane departure was imminent. However, in the absence of a model, the thresholds for 'too large' or 'too small' were determined empirically.

It could be argued that, to some extent, the adaptive road template matching work of Pomerleau [13] encapsulates our model in their learning network. However, since our model is derived analytically, there is no need for training.

2 Geometric model

Similar to other authors, Goldbeck and Huertgen [6] use a low order polynomial to model a curved road surface. Conversely, many authors, such as Lourakis and Orphanoudakis [9], find it acceptable to assume a flat road surface. The flat road model presents the opportunity to approximately invert the perspective effects caused by the projection onto the image plane.

Figure 1 depicts an elementary road model. The road is assumed to be locally flat, and the lane markings and boundaries are assumed to be locally parallel. The camera is positioned a lateral distance of Ω_+ from the lane marking and points down at the road with a tilt angle of α. The real world information is projected onto the image plane in accordance with the pinhole camera model. The dotted line in the image plane represents the projected points of the lane marking onto the image plane. The angle between this dotted line and the horizontal axis of the image plane is φ.

2.1 Lane angles and lane position

A yaw angle of β is introduced in Figure 2. Again, the dotted line comprises the projected points of the lane marking onto the image plane. The angle between the projected marking and the horizontal axis of the image plane is denoted as θ. Clearly,

$$\Omega_\beta = \Omega_+ \cos\beta.$$

But, considering the right hand triangle in the image plane, we also have

$$\Omega_\beta = \frac{h}{\cos\alpha \tan\theta}.$$

Hence, the lateral distance of the camera from the lane marking is

$$\Omega_+ = \frac{h}{\cos\beta \cos\alpha \tan\theta}. \tag{1}$$

This argument also holds with respect to the left-hand-side lane marking or boundary. Denote θ_- as the left lane angle and θ_+ as the right lane angle in the image plane. Furthermore, let Ω_-, and Ω_+ denote the lateral distance with respect to the left and right lanes, respectively. The relative distance of the camera with respect to the left lane, namely

$$\Omega^* := \frac{\Omega_-}{\Omega_- + \Omega_+},$$

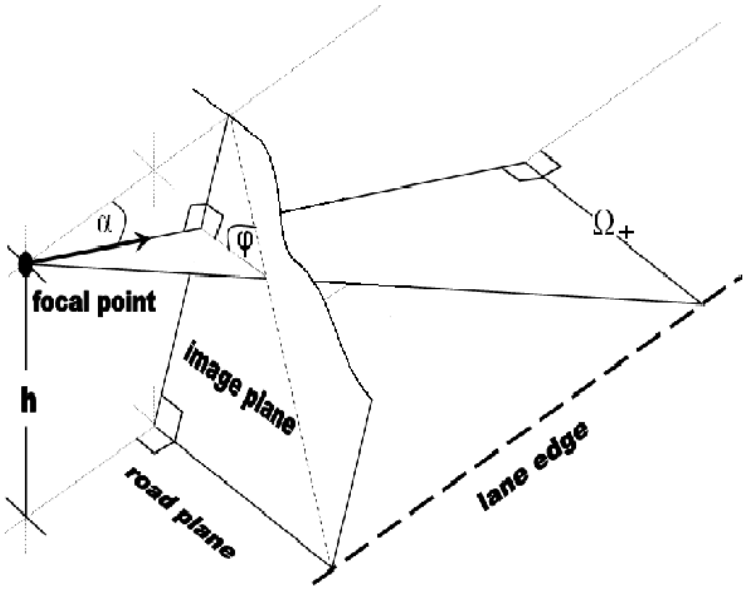

Figure 1: The road model. The tilt angle of the camera is denoted by α, the lane position by Ω_+, and the lane angle in the image plane by φ.

can therefore be expressed independently of the tilt and yaw angle, thus:

$$\Omega^* = \frac{\tan\theta_+}{\tan\theta_- + \tan\theta_+}. \tag{2}$$

The motivation for using relative, rather than absolute, distance comes from the application to driver behaviour studies. The lateral control of a car is more critical in a narrow lane, than in a wide lane. Drivers often take more care of their lateral position on small rural roads than on the wider lanes of motorways.

A contour plot of the function Ω^* is plotted with respect to the two lane angles θ_- and θ_+ in Figure 2.1. By inspection, it is clear that, should the lane angles both be small, or both large, the relative lane position function Ω^* will suffer from low error tolerance with respect to the measurements of θ_- and θ_+.

2.2 Optimisation

Of course, although the two lane angles are related to Ω^*, they are also constrained by the height of the camera h, the lane width Ω, the tilt angle α, and the yaw angle β, thus

$$\theta_- = \arctan\frac{\gamma}{\Omega^*} \tag{3}$$

$$\theta_+ = \arctan\frac{\gamma}{1-\Omega^*}, \tag{4}$$

with

$$\gamma := \frac{h}{\Omega\cos\alpha\cos\beta}.$$

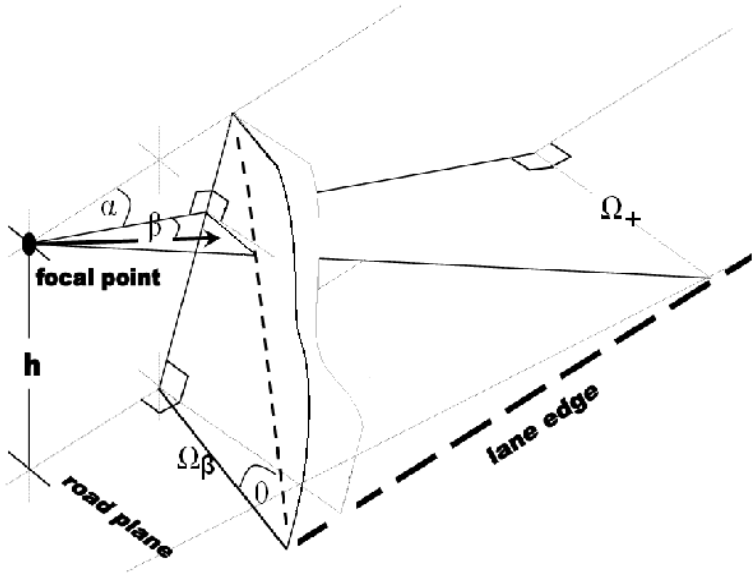

Figure 2: The generalised road model, which includes the yaw angle β. The tilt angle of the camera is denoted by α, the lane position by Ω_+, and the lane angle in the image plane by θ.

Using a few values of γ, Figure 4 depicts the curves (θ_-, θ_+), i.e. the curves

$$\theta_- = \arctan \frac{\gamma \tan \theta_+}{\tan \theta_+ - \gamma}, \tag{5}$$

plotted over the lane position function, subject to the constraints (3) and (4). This illustrates that, as γ tends to either zero or infinity, small errors in the measurements of either θ_- or θ_+ will have relatively large changes with respect to the estimated lane position. To find the values of (θ_-, θ_+) such that Ω^* is most stable we solve

$$\frac{\partial \Omega^*}{\partial \theta_-} + \frac{\partial \Omega^*}{\partial \theta_+} = 0.$$

By simple differentiation, the analytical solution is $\theta_- = \pi/2 - \theta_+$, that is when

$$\tan \theta_- = \frac{1}{\tan \theta_+}. \tag{6}$$

Substituting this result into (3) and (4) implies that

$$\gamma = \sqrt{\Omega^*(1 - \Omega^*)}. \tag{7}$$

If the height, yaw angle, and tilt angle of the camera are all fixed, this solution cannot be realised unless it is possible, by some electro-mechanical means, to alter the height, yaw or tilt, of the camera in real-time. Alternatively, if more than one forward looking camera is available, then one could set up the cameras to cover a range of fixed values for γ. Depending on the current value of the lane position Ω^*, the system could be optimised by taking the measurements of (θ_-, θ_+) from

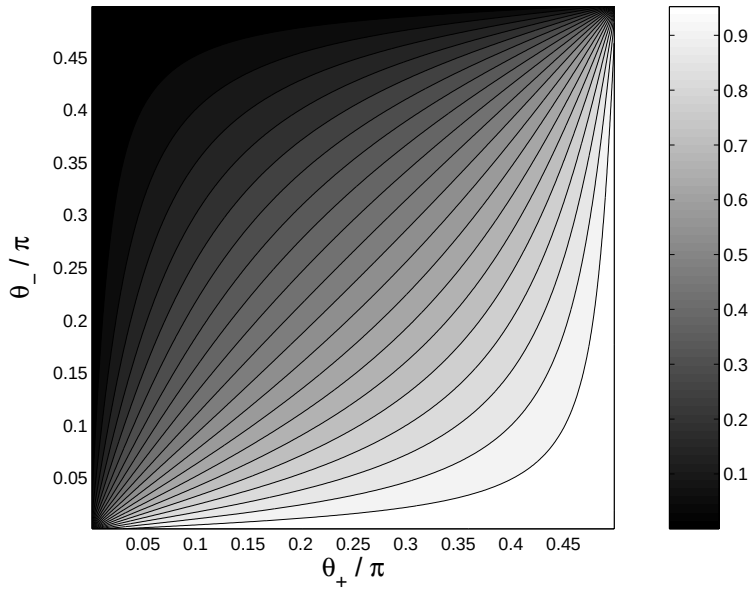

Figure 3: The relative lane position, represented by a shaded value from zero to one, is a function of the left lane angle θ_- and the right lane angle θ_+.

the camera with the nearest optimal value of γ. Another possibility is available in a multiple lane scenario. If the lanes are all of equal width, then the lateral position of the car can be measured with respect to the current lane or with respect to the total road width. This choice would alter the value of Ω and thus the value of γ. Again, the measurements of (θ_-, θ_+) could be taken from the lanes which yield the nearest optimal value of γ. Obviously, it goes without saying that such arguments are constrained to values of γ such that at least one left, and one right, lane marking is visible in the field of view. Such constraints will impinge upon other specifications, such as the internal parameters of the camera(s).

If γ is to be fixed, then it is still possible to find an optimisation with respect to solution (6). A natural approximation would be to minimise the mean squared error between (5) and $\theta_- = \pi/2 - \theta_+$, with respect to γ. That is, we seek $\operatorname{argmin}_\gamma D(\gamma)$, with

$$D(\gamma) := \frac{1}{\operatorname{mes}\Theta} \int_\Theta \left| \arctan \frac{\gamma \tan\theta}{\tan\theta - \gamma} - \pi/2 + \theta \right|^2 d\theta,$$

over the interval $\Theta = [\arctan\gamma, \pi/2]$. A plot of $D(\gamma)$ is shown in Figure 5. The solution, found numerically, is $\operatorname{argmin}_\gamma D(\gamma) = 0.44956062$.

Assuming that one does not have control over the road width, one is therefore left with the problem of finding $0 \leq \alpha, \beta \leq \pi/2$ and h such that

$$\frac{h}{\Omega \cos\alpha \cos\beta} \approx .44956062.$$

Again, the considerations of multiple cameras and/or multiple lane markings, as discussed above, can be used with respect to this result.

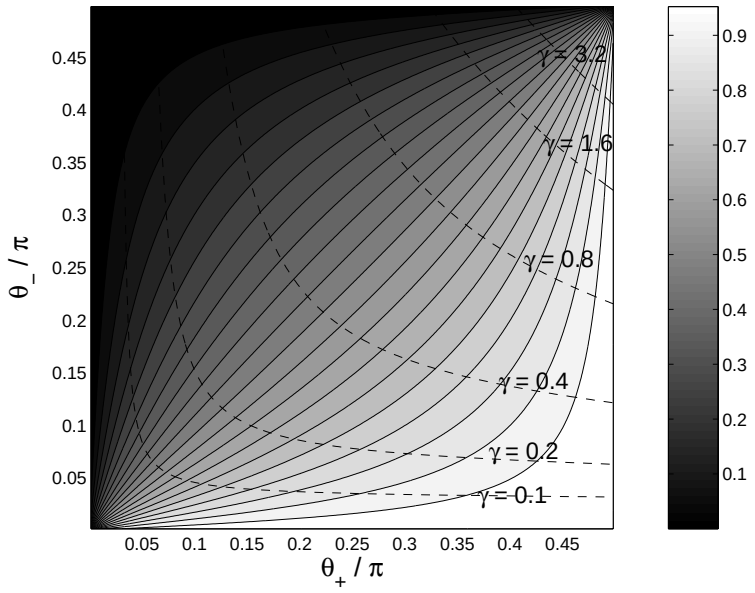

Figure 4: The relative lane position, represented by a shaded value from zero to one, is a function of the left lane angle θ_- and the right lane angle θ_+. The dotted contours depict the constraint governed by the parameter γ.

3 Lane detection and angle estimation

The task of lane detection has, itself, become a sub-field of computer vision. Developed in the latter part of the 1980's, the 'Personal Vehicle System' [7] described by Tsugawa [20] concentrates solely upon the lane detection problem. Once captured by a camera, the scene is convolved with the 3-by-3 Prewitt edge detection operator. The resulting image is then thresholded. Ideally, a white lane marking segment will yield two edges to this analysis. Any missed segments are sought by linear interpolation. From the experiments carried out, a success rate of 100% is quoted in day-time fair weather and 70% for night-time.

Another system surveyed by Tsugawa [20] is the 'Automated Highway Vehicle System' which was developed by Suzuki, Aoki, Tachibana, *et al* [18] [19]. Here the lane detection is more robust than the other systems he discusses. Before thresholding, the 3-by-3 Sobel edge detection operator is applied to the image. The remaining edges are kept if they are of a certain length and if the curvature is not too steep. In experiments, the final step was shown to eliminate unwanted edges such as shadows and tyre marks.

Indeed, the more robust lane detection algorithms must overcome common problems caused by dirt, shadows, and myriad environmental affects. These extraneous phenomena can partially obscure lane markings or cause false detection.

Because features such as dirt or shadows have localised effects it is not always enough to simply apply global methods like those surveyed by Tsugawa [20]. Instead a more sophisticated approach, resistant to local spurious phenomenon, has a greater likelihood of success.

Bertozzi and Broggi [1, 2] begin their lane detection procedure with an inverse perspective mapping. The lane markings are assumed to consist of light longitudinal

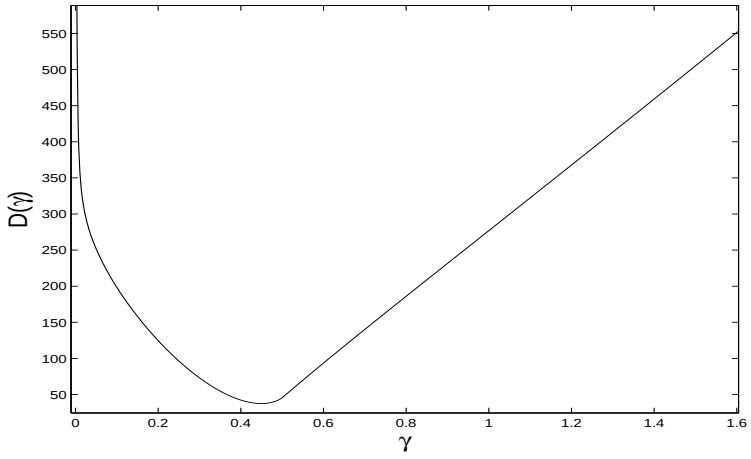

Figure 5: The mean squared error, with respect to the parameter γ.

regions surrounded by the darker regions of the road. Candidates with appropriately wide segments are passed through an iterative geodesic, morphological algorithm. After several iterations the procedure accentuated well connected edges. Other spurious edges were eliminated by thresholding. A local threshold is chosen which is proportional to the maximum value inside each neighbourhood of the image.

Charbonnier et al [3] developed a road marking detection method for an automatic repainting module. Assuming that the road marking width is known a priori they employed localised thresholding and a vertical edge detector. This was followed by a calculation of a series of local Hough transforms to 'account for the longitudinal coherence of road markings'. Given a suitably thresholded edge map, each Hough transform is taken by calculating a contour integral over a line, parameterized by angle and distance from the origin. A large value is registered when an edge coincides with the line of integration. That is, a linear model is fitted to the lane markings.

Jung and Kelber [8] employ a linear parametric method in the near field. Unlike Charbonnier, they model any markings in the far field with a quadratic, thus admitting the possibility of road curvature.

We follow a simplified version of Charbonnier et al [3] to detect the lane markings. A gradient operator ∇ is applied to the image $f : X \mapsto \mathbb{R}$, with $X \subset \mathbb{R}^2$, before the absolute values are thresholded via the operator

$$(\tau f)(\mathbf{x}) = \begin{cases} 1, & \text{if } f(\mathbf{x}) > f_p \\ 0, & \text{otherwise} \end{cases} .$$

In our system, the Sobel filter was used to approximate ∇; the threshold value f_p, was chosen semi-adaptively with respect to a predetermined parameter p, via

$$f_p = \frac{\|f\|_\infty}{\|f\|_1} \int\limits_{f<p} f d\mathbf{x}.$$

That is, a percentage $(1-p)$, of the pixels are expected to be edges. Assisted by a simple graphical user interface, the choice of p can be made empirically. Of course,

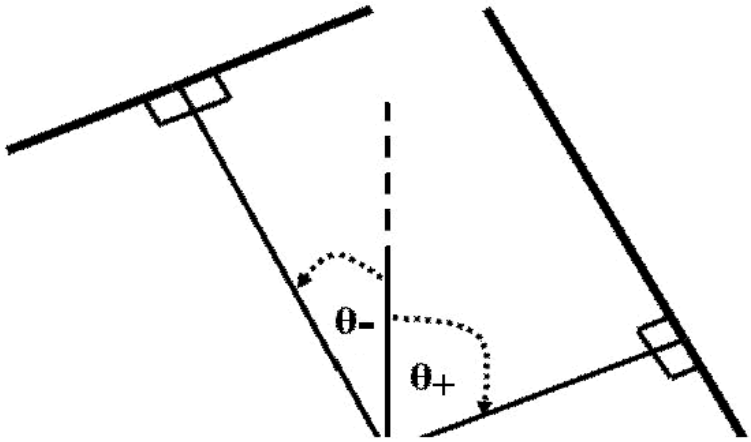

Figure 6: The two lane angles θ_- and θ_+ are found via two local Hough transforms.

the problem of choosing a threshold parameter can be circumvented by omitting τ, for then (8) would become a Radon transform.

Two local Hough transforms are performed on this binary map. The maximum of each Hough transform yields the lane boundary angles, viz.

$$\theta_\pm = \underset{\theta}{\operatorname{argmax}} \int_{\Lambda_\pm(\rho,\theta)} (\tau|\nabla f|)(\mathbf{x}) d\mathbf{x}, \tag{8}$$

with

$$\Lambda(\rho,\theta) = \left\{(x,\xi) \in X \subset \mathbb{R}^2 : \rho = x\cos\theta + \xi\sin\theta\right\},$$

and

$$\Lambda_- = \Lambda|_{\theta\in[-\pi/2,0]}\,, \qquad \Lambda_+ = \Lambda|_{\theta\in[0,\pi/2]}\,.$$

The two Hough transformations are illustrated in Figure 6. The origin $(0,0) \in X$ is located at the centre of the abscissa, and the bottom of the ordinates. Although rudimentary, this lane detection algorithm performed adequately to demonstrate the overall thrust of our methodology.

4 Experiments and results

Generally, qualitative arguments constitute the de facto validation technique for lane detection systems. The method is applied to a sequence of images and the success is judged manually. However, the lack of any 'ground truth' can only temper any conclusions drawn. Although McCall and Trivedi [12] also make this point, they established a approximate ground truth via a downward-pointing camera.

Since it is easier to manually judge the success of a lane detection method than a lane position measure, we have tackled the validation problem in two steps. Firstly, given a set of easily identifiable lane markings, the accuracy of the lateral position estimate is tested against physical measurements. Secondly, given a typical driving sequence, the success of the lane detector is validated manually.

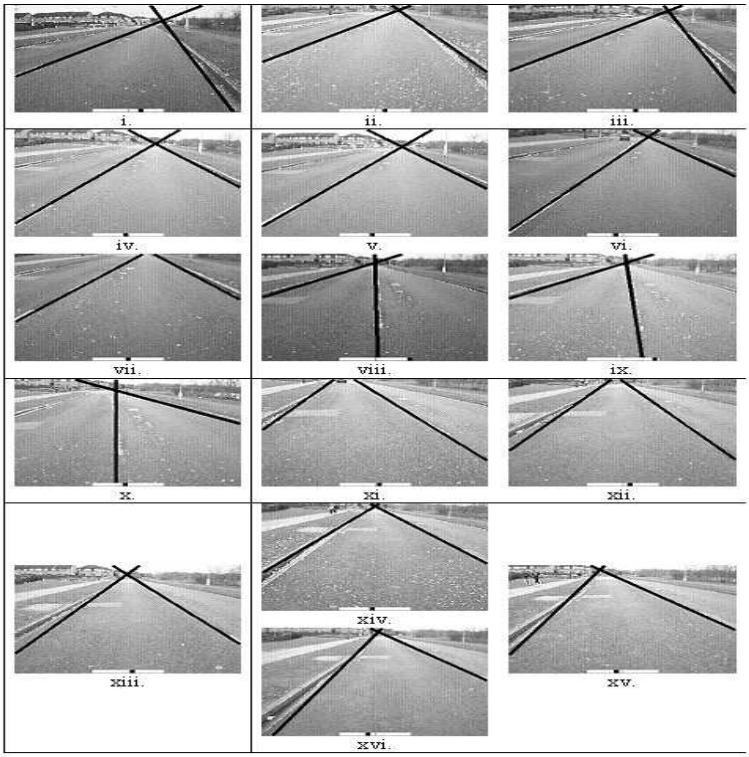

Figure 7: Image output of the lateral position gauge. The black lines indicate the estimated lane boundary.

The first validation step was carried out on a suburban road. A road was physically measured and marked with chalk. Some 16 images were captured of the road at various known lateral positions and unknown camera tilt and yaw angles. The results of the lateral position gauge could then be compared directly with the ground truth. The results, illustrated in Figure 7 and Table 1, indicate that, given the lane boundary detection method is somewhat successful, the lateral position estimate should result in an average error of some 4% of the lane width.

An image sequence of 4385 frames, taken at approximately 25 frames per second was processed by the lateral position gauge. The lane boundaries were successfully detected in 86% of the frames. Real-time tests suggested that the lane position algorithm performed between 12 and 15 Hz on a Pentium 1.6 GHz machine, depending upon the complexity of the image scene.

5 Conclusion

A simple model has been presented that relates the angles of two lane markings in the image plane with the lateral position of the host vehicle. Consequently, it has been shown that the relative position of the vehicle can be determined by the two lane angles alone, independent of camera height, yaw, and tilt.

Table 1: Quantative validation of the lateral position gauge

Frame	True Position	Estimated Position	Error %
i.	83	89	6
ii.	83	82	1
iii.	83	88	5
iv.	67	74	7
v.	67	73	6
vi.	67	70	3
vii.	75	75	0
viii.	50	50	0
ix.	50	48	2
x.	50	50	0
xi.	33	23	10
xii.	33	25	8
xiii.	33	23	10
xiv.	16	23	6
xv.	16	16	1
xvi.	16	16	1
Average			4.13

Furthermore, our model suggests an optimally stable set of values for the measurements of height, yaw, and tilt with respect to the relative lateral position.

Based on the Hough transform, a lane detection method was constructed to validate the theory.

References

[1] M. Bertozzi and A. Broggi. GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing*, 7:62–81, 1998.

[2] A. Broggi. Robust real-time lane and road detection in critical shadow conditions. *Proc IEEE Int. Symp. on Computer Vision*, 1995.

[3] P. Charbonnier, F. Diebolt, Y. Guillard, and F. Peyret. Road marking recognition using image processing. *International Conference on applications of leading-edge technologies to ITS. Boston*, 1997.

[4] E. D. Dickmanns and B. D. Mysliwetz. Recursive 3-d road and relative ego-state recognition. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14:199–213, 1992.

[5] L. Fletcher, N. Apostoloff, L. Petersson, and A.; Zelinsky. Vision in and out of vehicles. *IEEE Intelligent Systems*, 18(3):12–17, 2003.

[6] J. Goldbeck and B. Huertgen. Lane detection and tracking by video sensors. *IEEE International Conference on Intelligent Transportation Systems (ITSC 99). Tokyo, Japan*, 1999.

[7] A. Hattori, A. Hosaka, and M. Taniguchi. Driving control systems for an autonomous vehicle using multiple observed point information. In: *Proc. Intell. Vehicles*, pages 207–212. 1992.

[8] C. R. Jung, C. R. Kelber. A lane departure warning system based on a linear-parabolic lane model. *IEEE Intelligent Vehicles Symposium*, pages 891–895, 2004.

[9] M. Lourakis and S. Orphanoudakis. Visual detection of obstacles assuming a locally planar ground. *Proc. CVPR*, 2:527–534, 1998.

[10] Q. T. Luong, J. Weber, D. Koller, and J. Malik. An integrated stereo-based approach to automatic vehicle guidance. *Fifth International Congress on Computer Vision*, pages 52–57, 1995.

[11] J. McCall and M. M. Trivedi. Visual context capture and analysis for driver attention monitoring. *IEEE Intelligent Transportation Systems*, 2004.

[12] J. C. McCall and M. M. Trivedi. Performance evaluation of a vision based lane tracker designed for driver assistance systems. *Computer vision robotics research laboratory technical report, University of California*, 2004.

[13] D. A. Pomerleau. Ralph: Rapidly adapting lateral position handler. *IEEE Symposium on Intelligent Vehicles*, pages 506–511, 1995.

[14] D. A. Pomerleau and T. Jochem. Rapidly adapting machine vision for automated vehicle steering. *Expert Intelli. Systems Appl*, pages 19–27, 1996.

[15] R. Risack, P. Klausmann, W. Krüger, and W. Enklemann. Robust lane recognition embedded in a real-time driver assistance system. *IEEE International Congress on Intelligent Vehicles*, pages 35–40, 1998.

[16] M. Rossi, M. Aste, R. Cattoni, and B. Caprile. The IRST driver's assistance system. *Technical Report 9611-01, Instituto per la Ricerca Scientificia e Technologica*, 1996.

[17] D. D. Salvucci. Inferring driver intent: A case study in lane-change detection. *Proceedings of the Human Factors Ergonomics Society 48th Annual Meeting*, 2004.

[18] T. Suzuki, K. Aoki, and A. Tachibana. An automated highway vehicle system using computer vision - a vehicle control method using a lane line

detection system. In: *JSAE Autumn Convention Proc.*, pages 157–160. 1992.

[19] T. Suzuki, K. Aoki, A. Tachibana, H. Moribe, and H. Inoue. An automated highway vehicle system using computer vision - recognition of white guidelines. In: *JSAE Autumn Convention Proc.*, pages 161–164. 1992.

[20] S. Tsugawa. Vision-based vehicles in Japan: machine vision systems and driving control systems. *IEEE Trans. Ind. Electr.*, 41(4):398–405, 1994.

[21] J. Woong Lee. A machine vision system for lane-departure detection. *Computer Vision and Image Understanding*, 86:52–78, 2002.

[22] J. Woong Lee, C. D. Kee, and U. K. Yi. A new approach for lane departure identification. *IEEE Intelligent Vehicles Symposium*, pages 100–105, 2003.