Adaptive Road Parameter Estimation in Monocular Image Sequences

Włodzimierz Kasprzak Heinrich Niemann
Dirk Wetzel

Bavarian Research Center for Knowledge Based Systems
Knowledge Processing Research Group
Am Weichselgarten 7, D–91058 Erlangen
E–mail: *name*@forwiss.uni-erlangen.de

Abstract

An approach to road parameter recognition in monocular image sequences under egomotion is described. Two stabilization mechanisms are integrated into the *basic recursive estimation procedure*: a short sequence averaging of individual measurements and a scalar filter for long–term stabilization. Besides the *bottom–up measurements* a second type of road measurements is provided – *synthetic measurements* derived from model–based road stripe tracking.

1 Introduction

A specific application area of *dynamic computer vision* (based on image sequence analysis) is the autonomous vehicle guidance on roads ([1]). Two common tasks of such vision systems are: 1) road border following and 2) detection and tracking of obstacles. In this paper an extension of the first task to a *road recognition* problem and an adaptive solution of this problem are described. Besides the road border following the road class (i.e. lane number) is recognized and the following road parameters are estimated: width, camera orientation against the road plane, the relative egocar position on the road plane. Related work about road recognition is described, for example, in [2] (road border following) and [3] (vanishing point–based road curvature estimation),

The most frequent approach for the stabilization task is a *recursive filtering* of the state parameter values ([4], [5]). But the important problem is how to detect the road elements robustly, in order to supply the filter with measurement data, and how to make certainty judgements of the measurements. In this paper the robustness of measurements is achieved by using two types of the measurement data: bottom–up and synthetic (top–down) measurements. The judgement schemes are based on short–sequence variances of detected road elements (bottom–up data) or are long–sequence variances of road stripes (synthetic data). Related work about

*This work was partly supported by the "Deutsche Forschungsgemeinschaft", Bonn, Germany, and partly by the BMW AG., Munich, Germany. Only the authors are responsible for the content of this paper.

estimation mechanisms for road analysis is described, for example, in [6] (short sequence tracking of image points) and [7] (adaptive object estimation).

2 The road recognition module

The *road recognition* module is part of a road object tracking system under ego-motion that is depicted in Figure 1. The system consists of an application–independent module for image contour detection and 2–D motion estimation [8], the 2,5–D road recognition module (top part) and of a model–based module for object initialization and object tracking [9].

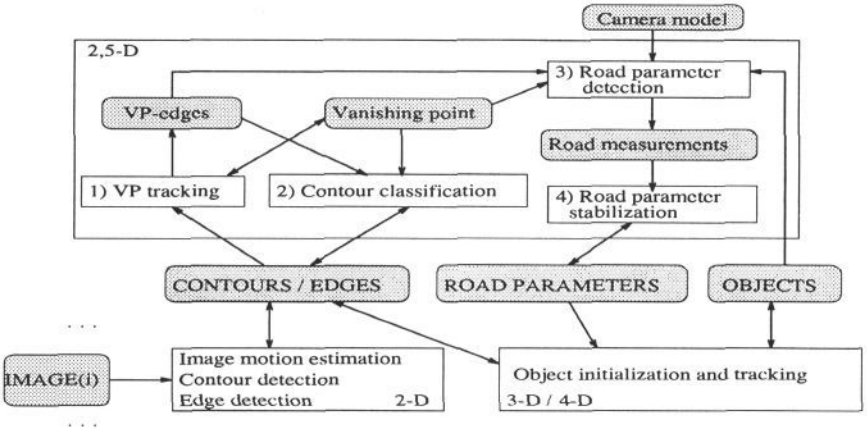


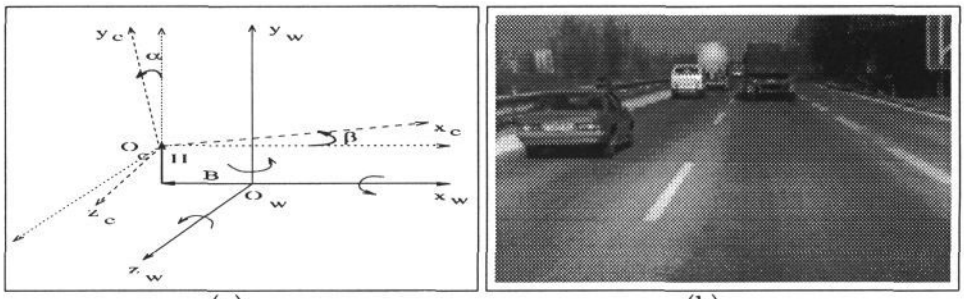
Figure 1: The road object tracking system structure, VP: vanishing point

The road estimation module interacts at one side with the 2–D module through image contours, supplies the 3–D object initialization unit with current road hypotheses and receives a feedback from the tracking unit in form of "synthetic" measurements.

Application–specific knowledge about the scene is used for the *vanishing point* (VP) detection and the detection of the *road border* and the *road stripes* in the image. Several road hypotheses could be generated which are characterized by different parameter sets. One road parameter set consists of the *road type* switch Ti, the road width W and the observer location relative to the road centre line B. The vanishing point and the road hypotheses are recursively stabilized during the image sequence processing cycle.

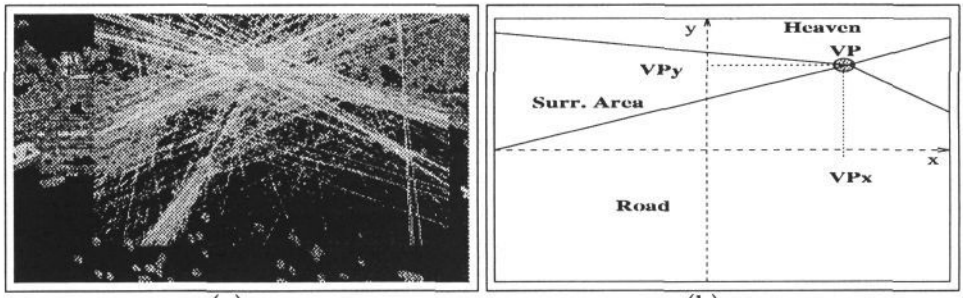
3 Road measurement detection

At time t_k of the k–th image in the sequence the transformation $EGO[H, \alpha(k), \beta(k)]$ of a point in the camera coordinate system (x_c, y_c, z_c) into a point in the road coordinate system (x_w, y_w, z_w) consists of a translation along y by H (the camera height above the road), a translation along x by $B(k)$ (the distance to the road centre), a rotation by $\alpha(k)$ around the x axis and a rotation by $\beta(k)$ around the



(a) (b)

Figure 2: The road–to–camera transformation (a) and one original image (b)



(a) (b)

Figure 3: Vanishing point detection (a) and the VP–based image areas for contour classification (b)

y axis (f_e is focus length) (Figure 2 (a)). The parameter H is assumed to be known and constant, but the remaining three transformation parameters have to be estimated.

3.1 Vanishing point tracking

The assumptions of approximate road planarity are made at this point of analysis. The non–horizontal edges of vertically elongated contours are supposed to be placed on the *VP–lines* (Figure 3(a)). The highest density area of intersection points between these lines is detected, and its centre point is the current *vanishing point measurement.* The vanishing point position is stabilized both in a short (up to 5 images) and a long sequence by a linear Kalman Filter (for details see section 4).

The rotation angles $\alpha(k)$ and $\beta(k)$ are directly derived from the location of the *vanishing point* in the image (Figure 3(b)):

$$\alpha(k) = -\arctan\left(\frac{VP_y(k)}{f_e}\right);\ \beta(k) = \arctan\left(\frac{VP_x(k)}{f_e}\right) \tag{1}$$

3.2 Contour classification

On the basis of current VP location the image pixels are classified into three classes: "road", "surrounding area" and "heaven" (Figure 3(b)). The contours containing some number of "road" pixels are classified as "road" contours, the contours without such pixels but containing enough "surrounding" area pixels are classified as surrounding contours. The remaining contours are assumed to be placed above the horizon.

3.3 Bottom–up road measurements

The VP–edges of "road" contours which are below the point VP in the image and are oriented towards this point are hypothesized to be the projections of road stripe borders. These VP–edges are backprojected into the road plane (due to the transformation $EGO(-H, -\alpha(k), -\beta(k)))$ (Figure 4). The weigths of back-projected edges are now determined, depending on the length, the distance from the observer, and the elongation along the z_w axis. An accumulator vector is provided, that corresponds to a cross section of the road (to a discrete segment of the Z^w axis) At the end the edges are projected onto the Z^w axis, and the edge weights are added in appropriate accumulator cells. The analysis of densities in this accumulator contours leads to different hypotheses about the road type – Ti (let us fix $i = 2, 3$) (where i is the lane number), the road width W_i and the location of the road grid relative to the observer (or dually the camera position B_i relative to the road centre line) (Figure 5). The certainty of each road type hypothesis Ti is evaluated also.

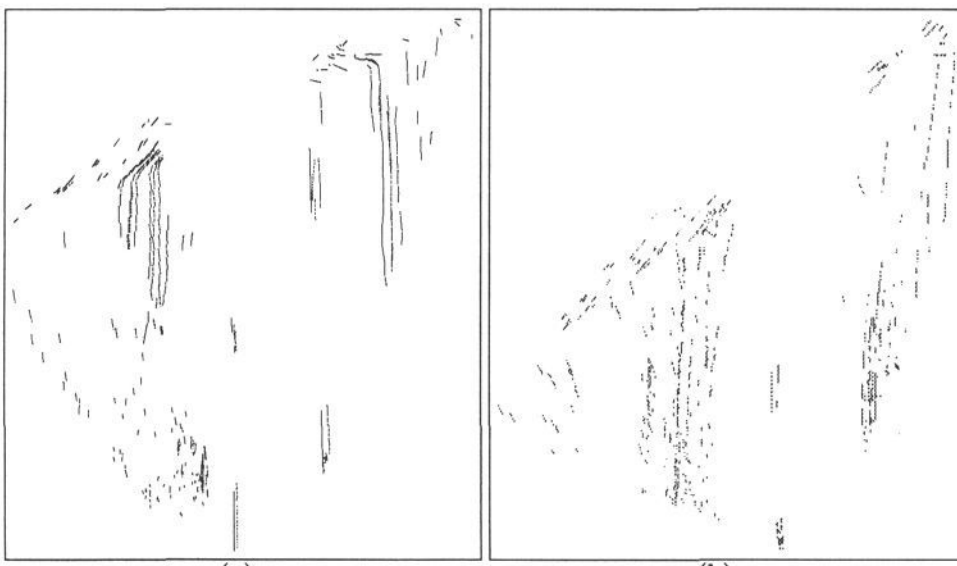
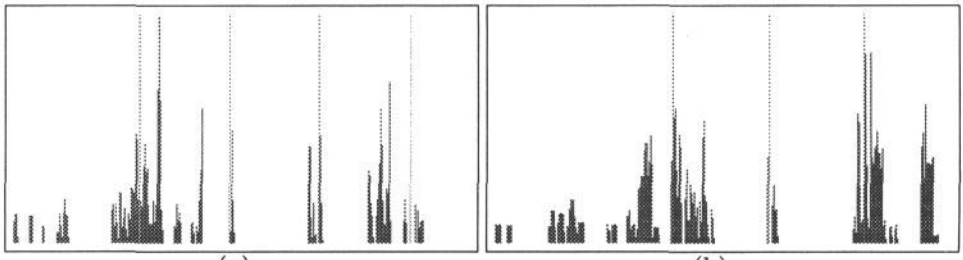


Figure 4: The backprojection of final VP–edges on the supposed road plane: (a) straight road, (b) curved road



(a) (b)

Figure 5: The weights in the accumulator vector corresponding to the final projection of the VP–edges onto the Z^w axis (histogram–like shape) and one road grid hypothesis: (a) a 3–lane hypothesis in a straight road case, (b) a 2-lane hypothesis in a curved road case

3.4 Top–down (synthetic) road measurements

The measurements detected directly in the image are not alone affecting the road parameter estimation. The egomotion detection procedure in the 4–D tracking module is looking at each new image for "stationary" road object hypotheses in order to estimate the *egomotion*. But the successfully tracked objects allow also a feedback to the road estimation module. Synthetic (top–down) measurements of the road parameters are extracted from the best "stationary" object set. Especially the selection of a proper road type hypothesis can now be improved. These top–down measurements induce a second road parameter stabilization step.

4 Road parameter stabilization

Let $\boldsymbol{s}_i(k) = [W_i(k), B_i(k)]^T$ be a 2–dimensional *parameter vector* at time t_k for the road type Ti $(i = 2, 3)$ and let $\boldsymbol{m}_i(k) = [w_i(k), b_i(k)]$ be an associated *measurement vector* containing the road measurements at time t_k. The time–dependent model of both vectors behaviour is given by a stochastically disturbed dynamic system with discrete time:

$$\boldsymbol{s}_i(k+1) = f[\boldsymbol{s}_i(k)] + \boldsymbol{v}(k); \quad (2)$$

$$\boldsymbol{m}_i(k) = h[\boldsymbol{s}_i(k)] + \boldsymbol{w}(k) \quad (3)$$

where $\boldsymbol{v}(k)$ is the system noise and $\boldsymbol{w}(k)$ the measurement error. The *filtering* task is to estimate the state $\boldsymbol{s}^*(t)$ on the basis of measurements $\boldsymbol{m}(t)$. A consecutive solution is achieved by *recursive* methods, where the old estimate of $\boldsymbol{s}(k)$ is updated after new measurements $\boldsymbol{m}(k+1)$ are available. The covariance matrices express the error probability of the measurement $\boldsymbol{m}(k)$ and of the state estimation $\boldsymbol{s}^*(k)$.

4.1 Weighted averaging of single measurements

The success of stabilizing each road parameter set $\boldsymbol{s}_i, (i = 2, 3)$ during adaptive filtering mainly depends from proper estimation of the measurement variances $\hat{R}_i$. Instead of working with measurements in one image a weighted average of up to

N ($N = 3-5$) corresponding measurements in the last N images is preferred here. The variances of individual parameter measurements in a short sequence of up to N images determine the current measurement variance of this parameter. Default minimum and maximum variances are provided, limiting the influence of single measurements on the estimated values.

The stabilization speed depends directly on the robustness of measurement detection and how far the initialization values are located from the real values. Besides the certainty estimation of measurements (which is opposit to the measurement variance) the weigthed averaging of corresponding measurements allows a more robust measurement detection (with subpixel accuracy) than in a single image case.

4.2 Adaptive estimation procedure

At time t_k of the k–th image there is a four–value set for each parameter $s_{ij}, (i = 2,3;\ j = 0,1)$ of a road hypothesis Ti : two measured values (detected and synthesized), and a predicted and modified estimation value. For the stabilization task a linear Kalman filter ([10]) is provided. The *recursive estimation* procedure for each road type hypothesis Ti consists of following steps:

1. *State vector initialization*: After the number k_0 of single measurements $m_j(k), k = 1, ..., k_0$ is greater than the number 2 of state variables, the first estimation $\boldsymbol{s}^*(k_0)$ and its covariance matrix $\boldsymbol{P}^*(k_0)$ are computed. A skip to step 5(c) follows.
2. *Bottom–up measurement*: At every new image $k > k_0$ a new measurement vector $\boldsymbol{m}(k)$ and its covariance matrix $\boldsymbol{R}(k)$ are determined. Thery are results of short–sequence weighted averaging of individual measurements.
3. First *stabilization* step: with (a) computation of the Kalman gain and (b) update equations for the estimated vector $\boldsymbol{s}^*(k)$ and its covariance matrix $\boldsymbol{P}^*(k)$ (but no prediction equation).
4. *Synthetic measurement*: On the basis of road stripe tracking a synthetic road parameter measurement $\boldsymbol{m}_S(k)$ and its covariance matrix $\boldsymbol{R}_S(k)$ are determined.
5. Second *stabilization* step: with (a) computation of the Kalman gain, (b) the update equations for estimations $\boldsymbol{s}^*(k)$ and variances $\boldsymbol{P}^*(k)$ and (c) the prediction equations for next time prediction of the parameter set $\boldsymbol{s}(k+1)$ and its covariance matrix $\boldsymbol{P}(k+1)$.
6. With $k \leftarrow k+1$ repeat from point 2.

4.3 Road class selection

The selection of the best road hypothesis is based on the relation between the variances P_2^* and P_3^* of the estimated road parameter B_2 and B_3, normalized over the road widths W_2 and W_3, respectively. The following rule can be applied for the selection of best road type hypothesis:

$$\text{IF}(P_3^* \leq \frac{W_3^2}{W_2^2}\ P_2^*) \text{ THEN } Class \leftarrow T_3 \text{ ELSE } Class \leftarrow T_2 \tag{4}$$

5 Results

The adaptive road recognition module was tested on several image sequences with 125 images of non–interlaced 350x282x8 bit resolution in each sequence. Images of motorway scenes and federal road scenes, with linear and curved road elements, have been provided.

In order to determine the detection and stabilization errors the original values have been measured in the image. For every image the original positions $\{(x_o, y_o)(k),\ (k = 0, ..., 124)\}$ of the vanishing point were manually determined. The road type Ti (2–lanes or 3–lanes) was obviously known. The real values of road width $W_o(k)$ and the original camera position $B_o(k)$ relative to the road centre line, have been manually measured in the backprojected images obtained during processing.

5.1 Vanishing point tracking

The results of detecting the vanishing point in 4 image sequences of real road scenes are summarized in Table 1. The second and fourth sequence are characterized by high road curvature. Nevertheless, while working with VP–edges which are relatively near to the observer, the original point was detected with good performance.

The differences between detected $(x(k), y(k))$ or stabilized values $(x^*(k), y^*(k))$ and the original values $(x_o(k), y_o(k))$ are represented by the error variables (X, Y) or (X^*, Yu^*). For every variable from the set $\{X, X^*, Y, Y^*\}$ its expected value $\mathcal{E}$ und variance σ^2 have been calculated.

The mean error values of VP detection and estimation are below 6 *pixels*. The error variances of the estimated values are from 5 to 10 times lower than the error variances of the measurements.

The stabilization speed depends directly on the relation between the measurement variances $R_{VP}(k)$ and the estimation variances $P^*_{VP}(k)$. The applied measurement variance estimation scheme, based on short sequence measurements, allows a stabilization of the estimations after 20 images with the uncertainty value in the covariance matrix $\boldsymbol{P}(k)$ limited to app. $40 \times 20[pixel^2]$.

Seq.	$X = x_{VP} - xo$		$X^* = x^*_{VP} - xo$		$Y = y_{VP} - yo$		$Y^* = y^*_{VP} - yo$	
	$\mathcal{E}_X$	σ^2_X	$\mathcal{E}_{X^*}$	$\sigma^2_{X^*}$	$\mathcal{E}_Y$	σ^2_Y	$\mathcal{E}_{Y^*}$	$\sigma^2_{Y^*}$
S 1	−1.57	66.82	−1.07	7.15	5.15	22.37	5.12	3.12
S 2	−5.18	17.17	−4.42	0.73	−0.82	6.91	−0.91	1.25
S 3	−3.67	80.35	−6.18	21.08	−3.06	17.06	−2.72	3.17
S 4	−3.51	155.64	−4.66	11.54	−6.37	63.69	−6.01	6.59

Table 1: Mean error and error variance of VP–detection in four image sequences

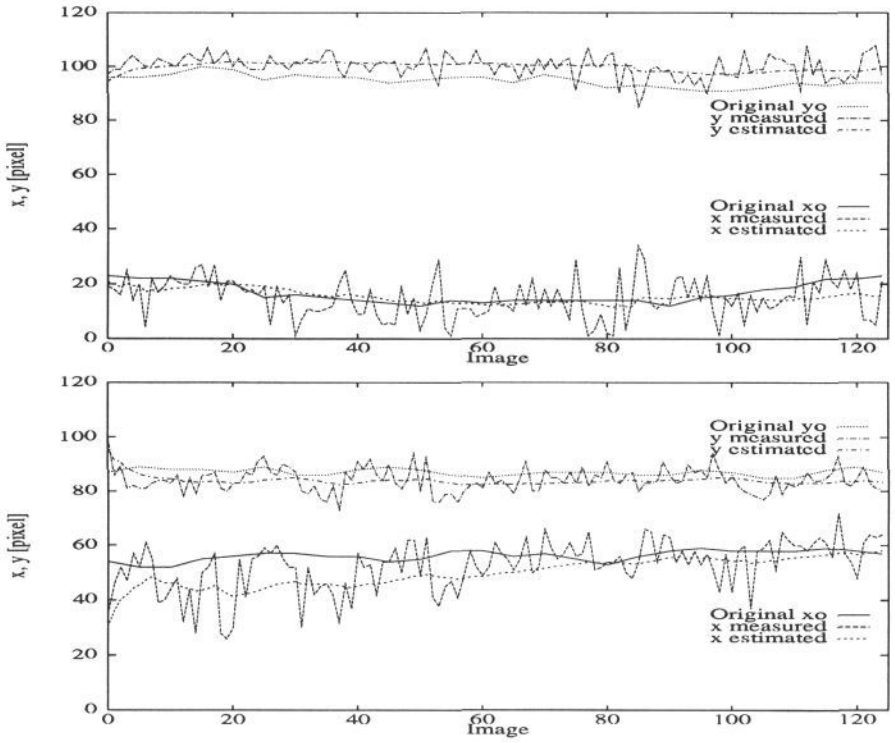


Figure 6: Vanishing point detection and stabilization in two sequences – with linear (top image) and curved (bottom image) road elements

5.2 Road parameter estimation

In Table 2 there are results presented from processing four image sequences by three stabilization methods: (a) single–image measurements, (b) short–sequence averaged measurements and (c) two stabilization steps. The estimated parameters from two road hypotheses $T2$ and $T3$ are related to the original road data. In particular, the stabilized values of road width – W_2^*, W_3^*, the observer location – B_2^*, B_3^* and the variances P_i^*–s of the estimated B_i–values are provided, as well as the number of images in which the given hypothesis has been selected.

The fine measurement detection procedures in case (b) and (c) allow a much better performance of the estimated road parameters for proper road type than in single stabilization of single–image measurements (case (a)). The variances and road parameter errors of proper hypotheses have been decreased by 50% while relating case (b) to case (a), and they have been decreased by another 20% if the case (c) is related to case (b). At the same time the variances of wrong hypotheses have been increased.

Seq. Method	$T2$: two lanes				$T3$: three lanes			
	B_2^* [m]	W_2^* [m]	$P^*{}_2$ [m^2]	Select $T2$	B_3^* [m]	W_3^* [m]	$P^*{}_3$ [m^2]	Select $T3$
1: (a)	-0.69	7.12	1.06	5	-0.82	10.52	1.85	118
1: (b)	-0.78	7.02	0.59	2	-0.83	10.54	0.84	121
1: (c)	-0.29	7.53	1.51	0	-0.78	11.14	0.52	123
	Original: B = -0.85 [m]; W = 11.00 [m]; Class = T3							
2: (a)	1.32	7.08	0.62	50	2.66	10.41	0.56	73
2: (b)	1.11	7.07	0.37	22	2.89	10.74	0.19	101
2: (c)	1.15	7.21	0.96	8	3.04	10.91	0.22	115
	Original: B = 3.00 [m]; W = 10.80 [m]; Class = T3							
3: (a)	1.31	7.28	0.54	101	0.84	9.68	1.21	22
3: (b)	1.39	7.66	0.11	112	0.62	9.47	1.01	11
3: (c)	1.62	7.47	0.10	117	0.42	10.42	0.92	6
	Original: B = 1.50 [m]; W = 7.10 [m]; Class = T2							
4: (a)	1.12	7.24	0.47	92	0.94	10.38	0.94	31
4: (b)	1.11	7.00	0.25	102	0.89	10.37	0.83	21
4: (c)	1.15	7.41	0.24	104	0.92	10.46	1.32	19
	Original: B = 1.35 [m]; W = 7.30 [m]; Class = T2							

Table 2: Applying three methods of road parameter estimation to four image sequences. The results for two road type hypotheses after processing 125 images

The columns *Select* T_2 and *Select* T_3 in Table 2 specify the number of images, for which the 2– or 3–lane road hypothesis have been selected. For stabilization method (a) the success ratio of proper road type selection is $60-90\%$. This ratio is increased in the case (b) to $82-98\%$ and in the case (c) it reaches $85-100\%$.

The stability of parameters for a proper road type is of satisfying performance and the mean error of B and W is below one meter. An optical evaluation of the tracking performance for W and B in two image sequences can be extracted from the Figure 7.

6 Conclusion

The characteristic features of proposed approach to road estimation are as follows:

- vanishing point lines are detected directly on the top of the camera vehicle $\Longrightarrow$ avoiding the problems with road curvature
- weigthed averaging of corresponding features in a short image sequence $\Longrightarrow$ more reliable estimation of image measurements than in single or two image case
- stabilization od road parameters $\Longrightarrow$ increasing the signal to noise ratio and filtering out the detection errors
- integrating the road detection with the model–based object tracking module $\Longrightarrow$ long–term road stripe tracking $\Longrightarrow$ synthetic road measurements are provided that improve the road recognition

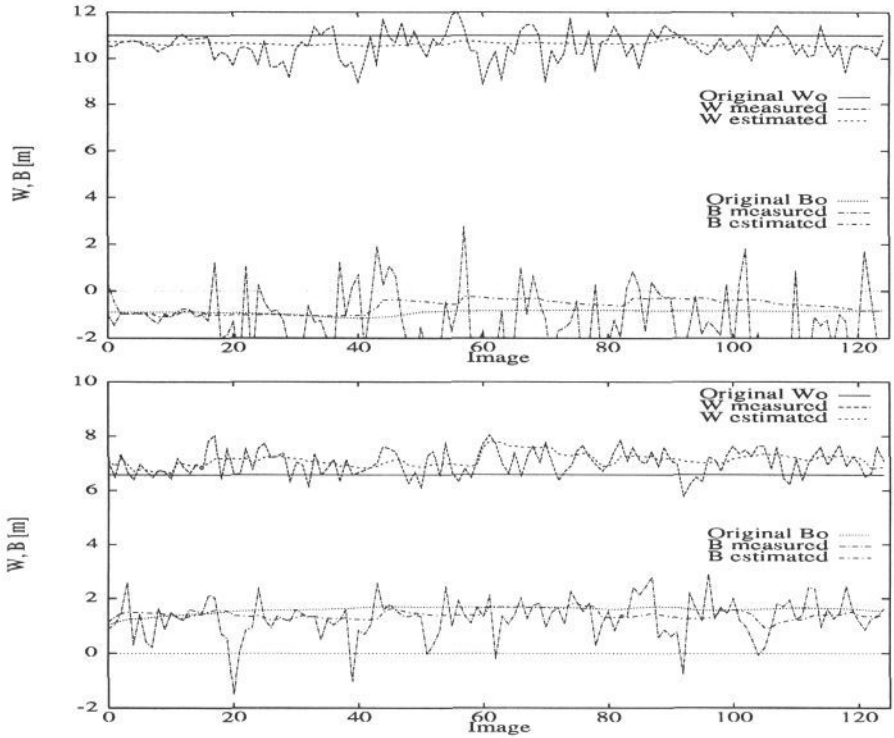


Figure 7: Detection and stabilization of road parameters in two image sequences: linear road (top image), (b) curved road (bottom image)

References

[1] Masaki I. (ed.): *Vision-based Vehicle Guidance*, Springer Series in Perception Engineering, Springer, New York Berlin Heidelberg etc., *1992.*

[2] Schaaser L.T., Thomas B.T.: *Finding Road Lane Boundaries for Vision-guided Vehicle Navigation*, In: [1], 238–254.

[3] Polk A., Jain R.: *A Parallel Architecture for Curvature-based Road Scene Classification*, In: [1], 284–299.

[4] Dickmanns E.D., Graefe V., *Applications of Dynamic Monocular Machine Vision, Machine Vision and Applications*, 1988, No. 1, 241–261.

[5] Gennery D.B.: *Visual tracking of known three-dimensional objects*, International Journal of Computer Vision, *7 (1992)*, 243–270.

[6] Salari V., Sethi I.K.: *Feature Point Correspondence in the Presence of Occlusion*; IEEE Transactions on Pattern Analysis & Machine Intelligence,*12 (1990)*, No.1, 87–91.

[7] Wu J.J., Rink R.E., Caelli T.M., Gourishankar, V.G.:, *Recovery of the 3-D Location and Motion of a Rigid Object Through Camera Image*, International Journal of Computer Vision, *3(1988)*,373–394.

[8] Kasprzak W., Niemann H., *Visual Motion Estimation from Image Contour Tracking*, In: D. Chetverikov, W.G. Kropatsch (Eds.), *Computer Analysis of Images and Patterns, Proceedings*, Springer, Lecture Notes in Computer Science, *vol. 719, 1993*, 363-370,

[9] Kasprzak W., *Road Object Tracking in Monocular Image Sequences Under Egomotion*, Machine Graphics & Vision, 3(1994), No. 1/2, ICS PAS Warsaw, 297–308.

[10] Jazwinski A.H., *Stochastic Processes and Filtering Theory*, Academic Press, New York and London, 1970.