Video-based Trailer Detection and Articulation Estimation

Lukas Caup[1], Jan Salmen[1], Ibro Muharemovic[2], and Sebastian Houben[1]

***Abstract*— Even for experienced drivers handling a roll trailer with a passenger car is a difficult and often tedious task. Moreover, the driver needs to keep track of the trailer's driving stability on unsteady roads. There are driver assistance systems that can simplify trajectory planning and observe the oscillation amplitude, but they require additional hardware.**

In this paper, we present a method for trailer detection and articulation angle measurement based on video data from a rear end wide-angle camera. It consists of two stages: to decide whether or not a trailer is coupled to the vehicle and to estimate its articulation angle. These calculations work on single video frames. The vehicle is therefore not required to be in motion. However, we stabilize the single frame estimations by temporal integration.

We perform training and parameter optimization and evaluate the accuracy of our approach by comparing the results to those of an articulation measurement unit attached to a test vehicle's hitch. Results show that it can very reliably be determined whether or not a trailer is coupled to the vehicle. Furthermore, its articulation can be estimated with a mean error of less than two degrees.

I. INTRODUCTION

Today's vehicles possess a variety of advanced driver assistance systems to provide comfort and security in difficult driving situations. These systems include traffic sign detection, emergency braking, lane control, blind spot assist and others. Assistance systems do exist for trailers, however, today they are deployed in trucks only since they require additional hardware. Nevertheless, most passenger car drivers do also consider handling a trailer as a difficult and cumbersome driving situation.

There are several situations that deserve special focus: Manoeuvring in narrow spaces, e. g. , in parking scenarios [1], requires careful handling of the vehicle to control the trailer's trajectory. This includes avoidance of jackknife where the trailer becomes uncontrollable [2] after exceeding a certain articulation angle. Furthermore, if the trailer starts to oscillate, especially on higher speed, a routine situation can quickly become dangerous [3].

In all these scenarios an accurate measurement of the trailer's articulation angle with respect to the vehicle is substantial in order to predict the upcoming trailer trajectory and pass that information to the driver. Additionally, the oscillations over time should be tracked to raise a warning if the amplitude or oscillation rate exceeds a given threshold.

In this paper, we propose to measure this angle with a rear-view fisheye camera. This hardware and some general computation unit can be found in today's medium-sized vehicles as they are used in situations like backup control or obstacle recognition during parking.

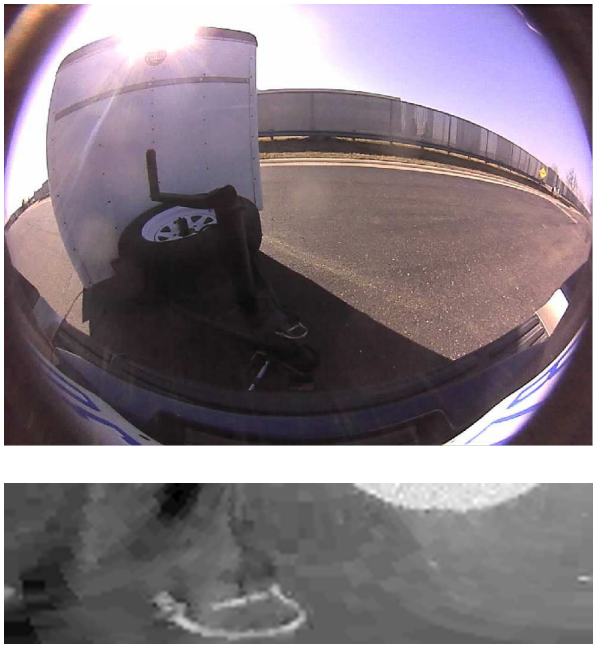


Fig. 1: An example image for articulation estimation. The trailer's shadow creates a symmetric structure (upper right corner) in the transformed image (logarithm shown) that may be confused with the drawbar axis, as it creates strong edges whereas the drawbar itself is hard to detect. This situation confirms the need for a relative brightness measure and a careful selection of all symmetry maxima.

We present a simple and fast video-based trailer assistance framework while relying solely on this built-in hardware aiming at middle and upper class vehicles that are occasionally used for trailer tours. Our approach makes very little assumptions on the trailer's visual appearance, namely, it only expects the drawbar to be symmetric. This rather general detection approach allows for an application on arbitrary trailer models with no need for training data.

II. RELATED WORK

Video-based driver assistance systems are increasingly important for intelligent vehicles. Today's systems already make use of many different camera systems (e. g., monocular/stereo, grayscale/Bayer-pattern, visible light/infrared, etc.) that can be mounted at different positions and with different perspectives. For reasons of pedestrian safety the

[1]Institute for Computational Neural Science, University of Bochum firstName.lastNameATini.rub.de

[2]Contintental Automotive Systems, Auburn Hills, Michigan ibro.muharemovicATcontinental-corporation.com

USA will soon introduce an obligation for rear-view cameras in every new vehicle [4].

The use of wide-angle (fisheye) cameras is especially promising since the observable surrounding is much larger with nearly identical hardware costs. Installing two (up to six) fisheye cameras, a complete view of the vehicle's surrounding can be provided eliminating blind zones for the driver, *cf.* [5][6][7]. For a detailed overview over technical aspects like distortion correction and transformations, we refer to [8].

Fisheye images available in a car can not only be displayed to the driver – but also computer vision algorithms can be employed in order to allow for advanced assistance applications. Wide angle cameras have been used for rear-view lane detection [9], for intersection assistance [10], and for obstacle detection [11][12], just to name a few. The strong distortion has to be accounted for in specialized image processing algorithms that usually are more demanding in term of computation costs than their pendants on narrow-angled cameras.

As soon as rear-view cameras are applied in a growing number, it stands to reason also to consider these for trailer monitoring. Taking into account the relatively short distance between the car's rear end and the trailer, wide angle cameras are particularly suited for this application.

To our best knowledge, there is no study yet considering video-based trailer detection and estimation of articulation angle. However, there are patents that cover contactless sensing in general for detection [13] and articulation estimation [14]. Another patent considering video-based articulation estimation has been applied [15] but remains unspecific on the image processing pipeline. The authors solely present a tracking algorithm for a trailer with an artificial pattern attached. In the paper at hand, no assumptions about the trailer appearance are made. Nonetheless, the algorithm parameters do implicitly contain information on a coarse scene representation (*e. g.* the position of the hitch, the maximum length of the drawbar within the image)

III. METHOD

The procedure consists of two subsequent steps: the detection stage determines whether or not a trailer is coupled to the vehicle, in that case the second step estimates the coupled trailer's articulation angle, and subsequently the articulation angle is integrated over time and the most plausible hypothesis is selected. The only information provided a priori is the image position of the hitch. We do not make any assumption about the trailer's appearance.

A. Detection of trailer

Knowing the hitch's position in the camera image the detection of a trailer can straightforwardly be solved by examining its surrounding (*cf.* Fig. 2a). To learn the appearance of asphalt and other ground structures we chose a subset of image sections around the hitch from sequences recorded without a trailer attached and clustered them to k centers by Kruskal's algorithm with the normalized cross-correlation coefficient (NCC) as distance function.

The resulting average image sections are prototypes for different pavement textures (*cf.* Fig. 4). If this very image section from an unknown frame is to be classified the highest of its NCCs with all the prototypes can be regarded as a likelihood of the absence of a trailer.

For the choice of an adequate training image set as well as the parameters k and a threshold τ for the resulting likelihood we refer to section IV.

B. Estimation of articulation angle

Since we do in general not have any information about the trailer's appearance, we decide to take advantage of the symmetric geometry of the trailer drawbar at least in proximity to the hitch. To simplify this task, we take an image section around the image position of the hitch C and calculate polar coordinates assigning a pixel (ϕ, r) with angle ϕ to the image row and distance r to C the gray value of the respective image pixel. The result is shown in figure 2b.

The symmetry search now reduces to a linewise search with an inspection of all columns as possible symmetry axis at the end.

Depending on weather conditions the trailer's drawbar often is shadowed by its setup. It is therefore imperative to apply a relative intensity difference model. We decided to take the natural logarithm of every pixel's intensity value mimicking the human relative brightness perception. The resulting image is processed by a gradient operator. We only take those pixels into account whose ϕ-component, i. e., the horizontal direction in figure 3a, exceeds the r-component, i. e., the vertical direction, by a given factor α. This favours edges that are aligned radial to the hitch. We found $\alpha = 2$ to work fine. A result of this step is shown in figure 3d.

The symmetry measure is now taken by a voting scheme that linewisely (containing pixels with equal radial distance to the hitch) regards all pairs of pixels with significant gradient magnitude and votes for their respective centre. These pixels are determined by linewisely selecting the n ones with the highest magnitude.

The vote image has the same size as the gradient image before and accumulates possible symmetry axes. An example is shown in figure 3e.

To determine the column with the dominant symmetry axis it is not sufficient to sum over all columns and subsequently take the maximum vote weight. A common source of false angle estimations is the presence of other symmetric structures receiving high symmetry votes as well. These structures can be distinguished from the drawbar's symmetry axis since their votes do not fill the whole column (refer to figure 1 for an example). However, depending on the magnitude of their edges the sum of all votes in that incorrect column can be of significant weight. We therefore calculate the p-quantile of all vote values in all columns to have a more robust estimator of the column's overall magnitude.

(a) (b)

Fig. 2: 2a The surrounding of the hitch that is taken into account for detecting the presence of a trailer. 2b The definition of the polar coordinate system to transform the area around the hitch for later symmetry analysis.

C. Temporal integration and plausibility check

To stabilize the angle measurements we apply an alpha-beta filter [16] that can predict future angle articulations by a simple constant velocity model which we feed back into the symmetry detection process by only looking for the next maximum in a vicinity of the predicted angle.

IV. EXPERIMENTS

We test our algorithm on 15 sequences with 2 different trailers recorded in Michigan, USA. The driving scenarios encompass slow manoeuvring with trailer articulation angles up to 50 degrees with respect to the driving direction as well as road and freeway situations at moderate velocity. The latter were recorded in natural traffic situations. Since weather and lighting conditions are always critical when doing real-world image processing, the test sequences also comprise both scenes with bright sunshine (and strong shadows) and cloudy weather. We therefore claim that the main areas of application for our system are covered.

The computations were performed on an Intel i7 at 2.00 GHz on a single core. We reached an average performance of 15 ms for the entire processing pipeline, trailer detection and articulation estimation.

A. Trailer detection

Table I gives an overview about the distribution of sequences to test, training, and prototype set. The naming is used throughout this subsection.

1) Clustering prototypes: To generate a representative collection of prototypes we choose a subset of three sequences without trailer extracting 150 image sections (PROTOTYPE SET). For the number of cluster centres k we examined $k \in \{1, 2, 5, 150\}$. Figure 4 shows the cluster centres of $k = 5$ clusters. We learn that it already contains clusters with only one example. Table II shows the best classification performance achieved on the training sets (*cf.* Sec. IV-A.2). Refer to figure 5 for a detailed analysis of the performance for different choices of the NCC threshold and k. Inspecting the false classifications also reveals that they do rarely occur in consecutive frames. It is therefore safe to say that the use of one or two cluster centres describes our prototype set best and yields an excellent classification performance.

TABLE I: Separation of training and test dataset for trailer detection performance

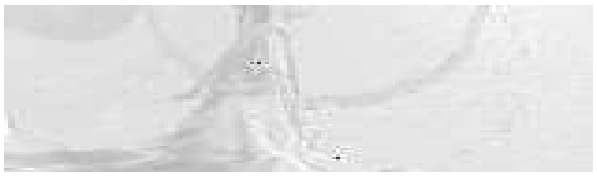
	#sequences (#frames) *without trailer*	#sequences (# frames) *with trailer*
PROTOTYPE SET	3 (150)	–
TRAINING A	1 (200)	1 (200)
TRAINING B	3 (600)	3 (600)
TEST	2 (200)	2 (200)

TABLE II: Performance regarding several number of prototypes

#Clusters Dataset	$k = 1$	$k = 2$	$k = 5$	$k = 150$
TRAINING A	99.3%	99.5%	92.5%	91.8%
TRAINING B	95.1%	96.1%	87.3%	82.1%

2) Parameter choice: For parameter choice we selected two training sets consisting of two sequences (TRAINING A, one with trailer, one without) and six sequences (TRAINING B, three with trailer, three without), respectively. None of these sequences coincide with the ones used for prototype clustering. By evaluating the algorithm trained on a small and a large training dataset we want to gather insight on how well the method generalizes to unknown situations and how stable the parameters can be set. Figure 5 shows the classification performance (trailer vs. no trailer) for different NCC thresholds and k. We can state that a NCC threshold from $[0.35, 0.5]$ combined with one or two representative

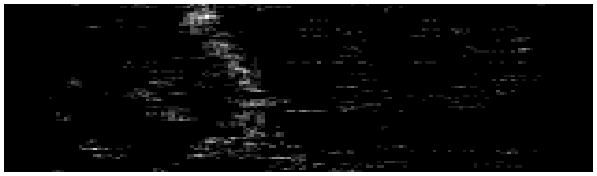
(a) Image section from figure 2b in polar coordinates



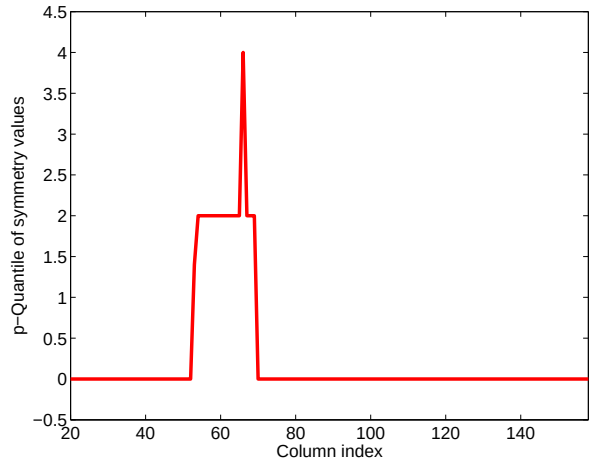
(b) Its logarithm

(c) Median filtering

(d) Salient gradients with high x component



(e) Symmetry measure on the gradients



(f) Columnwise p-quantile of above symmetry image. The column index of the maximum corresponds to the trailer's articulation angle.

Fig. 3: The processing line for estimating the articulation angle

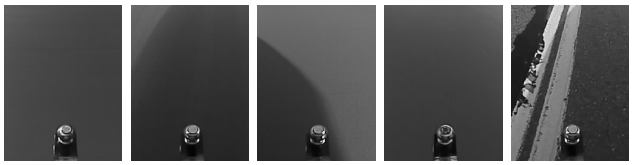


Fig. 4: Five of the prototype images that were clustered from hitch surrounding image sections. The cluster corresponding to the rightmost prototype did only consist of one example.

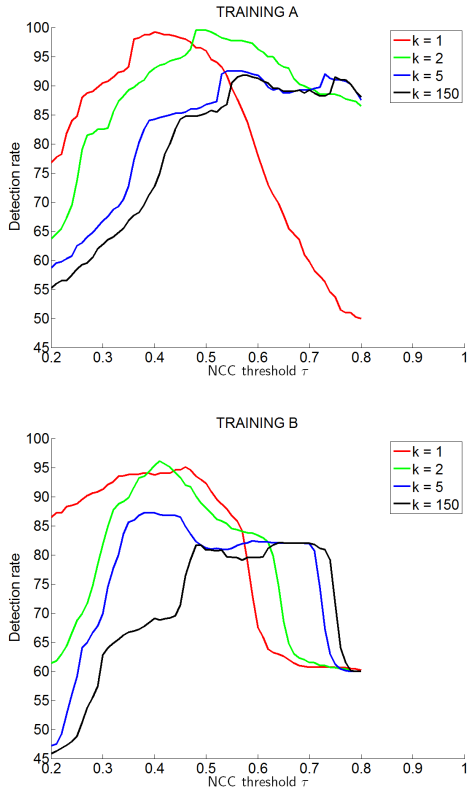


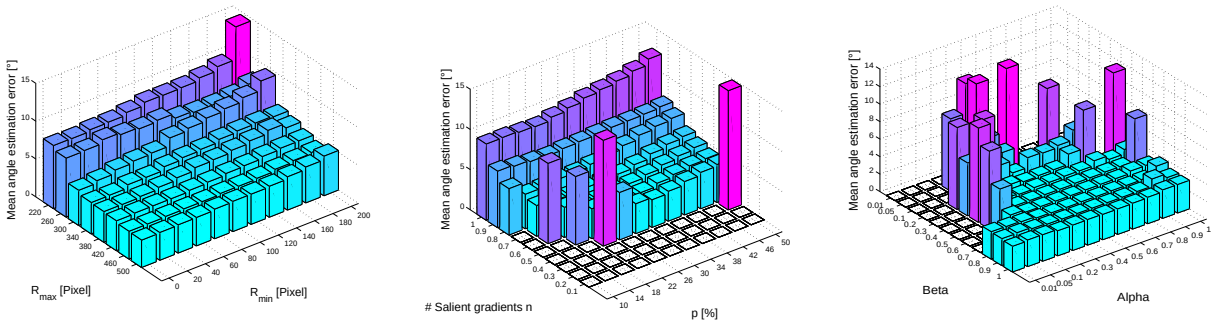
Fig. 5: Classification performance on TRAINING A and TRAINING B with varying NCC threshold τ for different number of clusters k.

prototypes is a good choice as it results independently from both training datasets. The course of the classification performance w. r. t. the NCC threshold is in both cases comparable which demonstrates the robustness and generality of our parameter choice.

B. Articulation estimation

To gather ground truth about the trailer's articulation we deployed a mechanical measurement unit at the hitch recording both video and articulation data. Since the unit was not available in all sequences we collected further ground truth by manually marking the drawbar in every video image. By this means we arrive at 4 sequences for training or parameter choice and 3 for testing.

The estimation procedure is parametrized with several quantities and optimizing them all is infeasible. However,



(a) (b) (c)

Fig. 6: Average error of angle estimation w. r. t. several parameters. Cases in which the drawbar could not be tracked are shown as empty bars. (6a) Average error of angle estimation w. r. t. the size of the considered hitch vicinity r_{min}, r_{max}. (6b) Average error of angle estimation w. r. t. the number of salient gradients n and the columnwise symmetry quantile p. (6c) Average error of angle estimation w. r. t. the parameters α, β of the temporal filter

we expect the parameters with the greatest influence on the estimation result to be

- the number of salient gradients n that vote for their respective symmetric centre
- the quantile p of symmetry votes that is calculated for each column in the vote image
- the circular surrounding defined by r_{min} and r_{max} of the hitch that is considered for symmetry measures (*cf.* Fig. 2b)
- the alpha-beta filter's noise and uncertainty parameters α and β

Other values like the sample density of the semi-circle do not play such an important role if not chosen too coarse. We evaluate the estimation error on the training set for different pairs of p, n, r_{min}, r_{max}, α, and β. The results are shown in figure 6.

Regarding r_{min} and r_{max} we conclude that our method performs best when examining a large vicinity around the hitch. As a matter of fact, the best values are achieved for $r_{max} = 500$ pixels which takes the whole drawbar into account. This is even more noteworthy as one would expect shadows to have a stronger negative influence on the measurement the larger the vicinity becomes.

Figure 6b reveals that p is a crucial parameter to the articulation estimation and does, as we expected, highly depend on the number of salient gradients. In fact, the choice of p affects how robust the measurement is against local symmetric structures (*cf.* Fig. 1).

The tests with varying α and β were performed with the optimal parameters r_{min}, r_{max}, n, and p that were computed beforehand.

V. RESULTS

A. Trailer classification

In view of the results of the parameter training in section IV we choose $k = 2$ prototypes, $\tau = 0.5$ via TRAINING A, and $\tau = 0.4$ via TRAINING B respectively. The test setup is composed of 4 sequences with 400 frames (200 with trailer, 200 without).

TABLE III: Classification performance in test scenarios

Sequence	Detection rate (no trailer)	Detection rate (trailer)
TEST A ($\tau = 0.5$)	76.5%	100.0%
TEST B ($\tau = 0.4$)	99.5%	95.0%

We learn that the parameters from TRAINING A are insufficiently precise to achieve a reasonable result during this test scenario (*cf.* Tab. III). Nonetheless, the results with the more complete training set TRAINING B are very promising. We conclude that τ is indeed crucial to the algorithm performance and parameter training should therefore be based on an adequately large data set. The strong model assumptions, $k = 2$ cluster centres, lead to a nearly identical performance between test and training data set as the latter is strongly generalized by the clustering algorithm. Closer inspection of incorrectly classified images in TEST B reveals that errors seldom occur on consecutive frames. We can therefore consider this problem solved.

B. Articulation estimation

Examining the findings from section IV-B we apply $r_{min} = 0$, $r_{max} = 500$, $n = 32$, $p = 0.6$, $\alpha = 0.5$, and $\beta = 0.5$ as parameter test settings. Table IV shows that the average absolute deviation on angle measurements is at about 1.95degrees. Figure 7 shows the temporal course of the GT angles and their respective estimation. Since our algorithm provides angles in image coordinates, we introduced a mapping between the measurements of the mechanical unit and manually labelled articulations by an affine function.

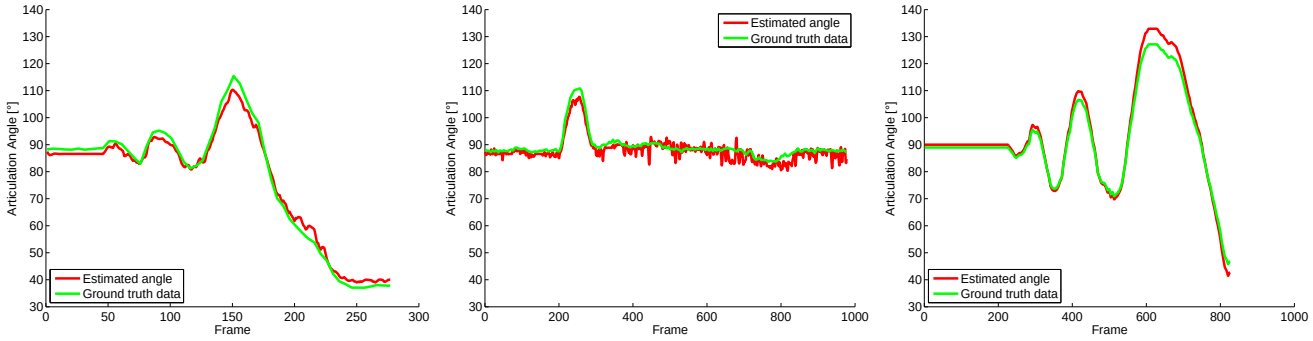


Fig. 7: Average error of angle estimation in test sequences over time.

TABLE IV: Overall angle estimation performance on test sequences. The asterisc marks the sequence wih ground truth determined by the mechanical measurement device.

	#Frames	Scenario	Avg. Estimation error [°] (maximum error)
TEST A	277	slow manoeuvring	2.19 ± 1.19 (5.48)
TEST B	977	countryside driving	1.69 ± 1.38 (7.55)
TEST C (*)	826	slow manoeuvring	1.97 ± 1.64 (5.78)
Total	2080		1.95

VI. DISCUSSION AND FUTURE WORK

In order to test the method's ability to predict trailer instability we want to perform longer test drives. Additionally, the use of other than the tested cameras should be considered in order to examine if the proximity of the hitch implies the need for wide-angle cameras or if a narrow-angle camera will suffice as these are currently installed more often in today's middle-class vehicles. Furthermore, different camera positions should be taken into account.

We were able to show that trailer monitoring with wide-angle cameras is a promising solution to provide trailer safety applications for middle-class vehicles with the necessary hardware that are occasionally used for trailer drives. Detection and tracking of the trailer's drawbar has been proven to be a reliable and general method. The parametrization of the proposed algorithm was shown to be stable with respect to different weather and lighting conditions as well as the attached trailer model.

REFERENCES

[1] D. Zöbel, "Minimum parking maneuvers for articulated vehicles with one-axle trailers," in *Proceedings of IEEE International Conference on Control, Automation, Robotics and Vision*, 2006, pp. 1–6.

[2] J. Morales, A. Mandow, J. L. Martínez, and A. J. García-Cerezo, "Driver assistance system for backward maneuvers in passive multi-trailer vehicles," in *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, 2012, pp. 4853–4858.

[3] P. Ridley and B. Stevenson, "Automated vehicle stability control for articulated vehicles," in *Proceedings of the Australasian Conference on Robotics and Automation*, 2005.

[4] "Pub. l. 110–189, 122 stat. 639: Cameron gulbransen kids transportation safety act of 2007," *110th Congress*.

[5] S. Li, "Monitoring around a vehicle by a spherical image sensor," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 4, pp. 541–550, 2006.

[6] T. Ehlgen, M. Thom, and M. Glaser, "Omnidirectional cameras as backing-up aid," in *IEEE International Conference on Computer Vision*, 2007, pp. 1–5.

[7] Y.-C. Liu, K.-Y. Lin, and Y.-S. Chen, "Bird's-eye view vision system for vehicle surrounding monitoring," in *Proceedings of the international conference on Robot Vision*, vol. 4931, 2008, pp. 207–218.

[8] C. Hughes, M. Glavin, E. Jones, and P. Denny, "Wide-angle camera technology for automotive applications: a review," *IEEE Intelligent Transport Systems*, vol. 3, no. 1, pp. 19–31, 2009.

[9] A. Takahashi, Y. Ninomiya, M. Ohta, M. Nishida, and M. Takayama, "Rear view lane detection by wide angle camera," in *Proceedings of the IEEE Intelligent Vehicle Symposium*, 2002, pp. 148–153.

[10] S. Gehrig, C. Rabe, and L. Krüger, "6D vision goes fisheye for intersection assistance," in *Proceedings of the Canadian Conference on Computer and Robot Vision*, 2008, pp. 34–41.

[11] M. Djalalov and T. D. Radjabov, "Top-view surrounding monitor system using modified-sift based stitching method with obstacle detection," *Advances in Electrical Engineering and Electrical Machines*, vol. 134, pp. 483–488, 2011.

[12] J. Molineros, S. Cheng, Y. Owechko, D. Levi, and W. Zhang, "Monocular rear-view obstacle detection using residual flow," in *Proceedings of the European Conference on Computer Vision, Workshops and Demonstrations*, 2012, pp. 504–514.

[13] S. J. Buckley, *US 7786849: Trailer Detection System*. Chrysler Group LLC, 2010.

[14] Y. H. Lee and A. Kade, *US 8073594: Trailer Articulation Estimation*. GM Global Technology Operations LLC, 2011.

[15] Y. Lu, J. S. Gibson, T. G. Johnson, and S. T. Hunt, *International Application No. PCT/US2012/022517: Rear Vision System with Trailer Angle Detection*. Magna Electronics Inc., 2012.

[16] E. Brookner, *Tracking and Kalman Filtering Made Easy*. John Wiley & Sons, Inc., 1998.