

# Investigating Camera Calibration Methods for Naturalistic Driving Studies

Jeffrey Paone,<sup>a</sup> Thomas Karnowski,<sup>b,\*</sup> Deniz Aykac,<sup>b</sup> Regina Ferrell,<sup>b</sup> Jim Goddard,<sup>b</sup> Austin Albright<sup>b</sup>

<sup>a</sup>Department of Computer Science, Colorado School of Mines, Golden, CO 80401

<sup>b</sup>Electrical and Electronics Systems Research Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

## Abstract

Naturalistic driving studies typically utilize a variety of sensors, including radar, kinematic sensors, and video cameras. While the main objective of such sensors is typically safety focused, with a goal of recording accidents and near accidents for later review, the instrumentation provides a valuable resource for a variety of transportation research. Some applications, however, require additional processing to improve the utility of the data. In this work, we describe a computer vision procedure for calibrating front view cameras for the Second Strategic Highway Research Project. A longitudinal stability study of the estimated parameters across a small sample set of cameras is presented along with a proposed procedure for calibrating a larger number of cameras from the study. A simple use case is presented as one example of the utility of this work. Finally, we discuss plans for calibrating the complete set of approximately 3000 cameras from this study.

**Keywords:** computer vision, naturalistic driving studies, camera calibration, sensor fusion

\* corresponding author's E-mail: karnowskitp@ornl.

## Introduction

Naturalistic driving studies (NDSs) consist of various recorded data from individuals driving in real-world environments. The Second Strategic Highway Research Project (SHRP-2), conducted between 2012 and 2013, included a NDS with a goal of addressing the role of driver behavior and performance in traffic safety.<sup>1</sup> The study included approximately 3,000 drivers over a 2-year period whose personal vehicles were specially equipped with a custom data acquisition system (DAS) designed by the Virginia Tech Transportation Institute (VTTI).<sup>2</sup> The DAS included kinematic sensors, a GPS, a forward radar, and video cameras which captured videos of the front, driver face, hands or dash area, and rear. The data has several privacy constraints which prevent open sharing.

The video data is recorded at 15 frames per second. The four videos are fused into a single frame, with some scaling and cropping used to achieve an overall size of  $720 \times 480$  pixels, as shown in Figure 1. The data stream is then compressed for efficient storage. Data may be requested for studies from VTTI through a formal process.<sup>3</sup> Typically the video is retrieved from the database as mp4 files that contain a single video camera stream. The data set is a tremendous resource for transportation studies, but the sheer size of the set would benefit from increased automation for processing.

One key step needed to improve the potential of the NDS to understanding driver behavior is camera calibration. Many applications in computer vision rely on the ability to make measurements from cameras, and camera calibration is an important component of many computer vision algorithms.<sup>4</sup> Extending the DAS camera from a recording device to a measurement device can yield a deeper understanding of vehicle dynamics and interactions.



**Figure 1.** Example of composite video frame. In this case the rear camera is not used, resulting in a blank image for the lower right pane.

Because the focus of NDS has been on safety, little attention in the literature has been given to other aspects of the data and processing,<sup>5</sup> especially outside the vehicle. However, there is a body of work on camera calibration that compares a camera self-calibration method against classical calibration, targeting vehicle-mounted applications (e.g., Reference 6), and explores multiple computer vision techniques (enhanced with calibration information) with a goal of providing real-time safety updates based on detected activity inside and outside of the vehicle (e.g., Reference 7). Multiple works also exist on the best practices to calibrate a single camera. Orekhov et al. present a calibration method that can be applied to any lens type.<sup>8</sup> Their distortion method follows a polynomial function, as opposed to the more common radial model. Brito et al. describe methods to automatically correct the radial distortion effects seen in images.<sup>9</sup>

While there are multiple methods to calibrate a single camera, there is no accepted metric on how to apply a single camera calibration model across several thousand cameras. With the rise of autonomous vehicles, much work in video processing for detection and analysis along with data fusion methods is moving from academia to actual implementations.<sup>10,11</sup> These methods all require accurate camera calibration.

In our work, we faced the special challenge of having to perform post hoc camera calibration, where data is obtained “in the wild” without concern for computer vision measurements. Consequently, we had the additional constraints of limited hardware access, an inability to modify installations, and other issues such as image quality.

This paper is organized as follows. First, we describe a laboratory experiment in which we calibrated a sample set of 50 SHRP2 camera units to determine variability and create a calibration model for the intrinsic parameters (encompassing distortion, focal

length, etc.). Next, we explore methods for extrinsic parameters which refer to the positioning of the camera relative to the roadway for our application. We give a sample application and conclude with a summary of our findings and recommended next steps.

## Intrinsic Camera Calibration

### Motivation

Computer vision applications in general relate collected imagery to real-world conditions that allow machines to respond, adapt, or interpret the world around them. Although we are dealing entirely with recorded video data in our application, our goal is to use video imagery to take measurements that will answer many questions in transportation research, particularly with respect to how drivers behave under different conditions. Essentially, we would like to transform the camera from a recording device to a true measurement device. The front video camera of the SHRP2 DAS was selected for this study, as it can provide a wealth of information regarding a driver's environment. The front video feed has an image size of  $480 \times 360$  pixels. In camera calibration, the first step is typically to estimate the intrinsic parameters, which include the functioning of the lens and the relationship between the lens and the imaging sensor. We use a nine-parameter model for the intrinsic parameters, which includes  $x$  and  $y$  components of focal length and optical center, and five distortion parameters. Various software tools to extract these parameters are available, which typically involve the use of a known calibration target with easily localized features such as checkerboard corners or circle centers.<sup>12</sup> Good overviews can be found in a variety of textbooks.<sup>13</sup>

### Experimental procedure

We borrowed 50 DAS units from VTTI and custom software to record data from the DAS. We also obtained a breakout electronics component that allowed us to better access each camera individually instead of the composite frame. Access to the cameras was achieved through a USB-NTSC interface, avoiding potential compression artifacts in the recorded stream.

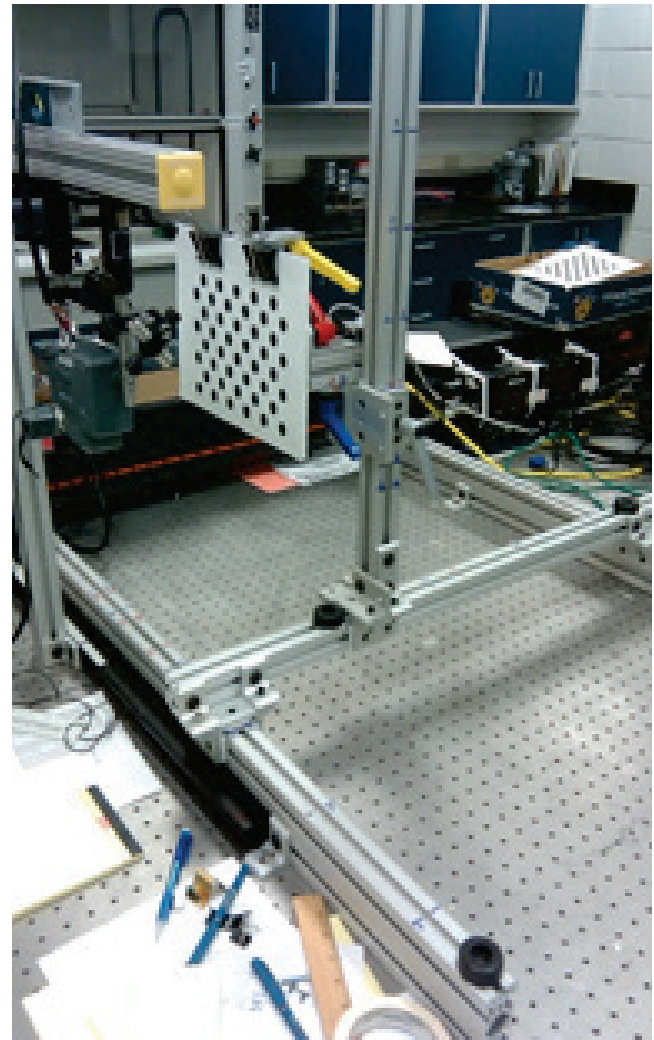
The full SHRP 2 study consisted of approximately 3,000 camera units. Our experimental calibration procedure involved a sample of 50 camera units. While the sample consisted of less than 2% of the full study, every camera unit had identical hardware, and any differences were attributed to imperfections in the individual camera units. Consequently, each camera was fully calibrated to understand the differences between the individual cameras.

We acquired data for the camera calibration using a laboratory procedure designed to present consistent views of the calibration target to the camera during testing. A laboratory rig with six degrees of freedom was built which allowed us to mount the DAS in a consistent position and pose. The laboratory calibration rig is shown in Figure 2, with the DAS head unit suspended on the left side of the image.

We then moved the calibration target through a series of poses using a system of rails and sliding mounts. We chose an asymmetric circle pattern as opposed to a checkerboard because the corner features in the latter were more challenging due to analog camera interlacing. We detected the asymmetric circular grid patterns in each frame using the OpenCV calibration routines. A typical sequence would generate 6,000 frames of data, with roughly 2,000 frames having a detected pattern.

For the frames in which the pattern was successfully detected, we applied motion stability criteria by computing the homography for the frame and then used only those frames that showed a minimal

change from the previous frames. A rich set of pose frames was then created by using the homography matrix and computing a dissimilarity metric. The metric was computed by finding the relative difference L2 matrix norm between the homography of pairs of frames. Any pair that exceeded a given threshold was retained as being sufficiently dissimilar examples of the target pose and position. The threshold was manually adjusted until roughly 80 frames per camera were achieved; in practice, the typical dissimilarity threshold was 0.1, with smaller values yielding more frames.



**Figure 2.** Calibration rig prepared for the 50 cameras. The asymmetric circle target is positioned in front of the SHRP2 camera unit.

We then computed the intrinsic parameters using an iterative process. After the first iteration of parameters was computed, the frames in the set were projected and a root mean square (RMS) error was computed. Frames in which the error exceeded the average error by a factor of 3 or more were removed from the calculation to eliminate outliers that would contribute to a poor calibration. We then repeated the computation until no frames had to be removed and the calibration was stable. As a final check, the undistorted

image was generated as well, warping all pixels of the image for a complete view of the rectification.

## Results

Once the camera calibration was completed, the parameters were used to reproject the grid's three-dimensional world coordinates to the image's two-dimensional pixel coordinates. The difference between the detected point and projected point was computed to determine the reprojection error.

In order to understand the performance of a given set of intrinsic parameters, the intrinsic parameters are applied to every

camera's video stream. For each camera, we computed the intrinsic and distortion parameters. The computed parameters from a single camera (e.g., Camera 01) were then applied to the other cameras (e.g., Cameras 02–50) to calculate the RMS reprojection error. These pairings were then used to calculate how similar two cameras were to one another. Based on this analysis, we determined that Camera 59 was most like the other cameras in the dataset and therefore was selected for the general model (Figure 3). The parameters are shown in Table 1.

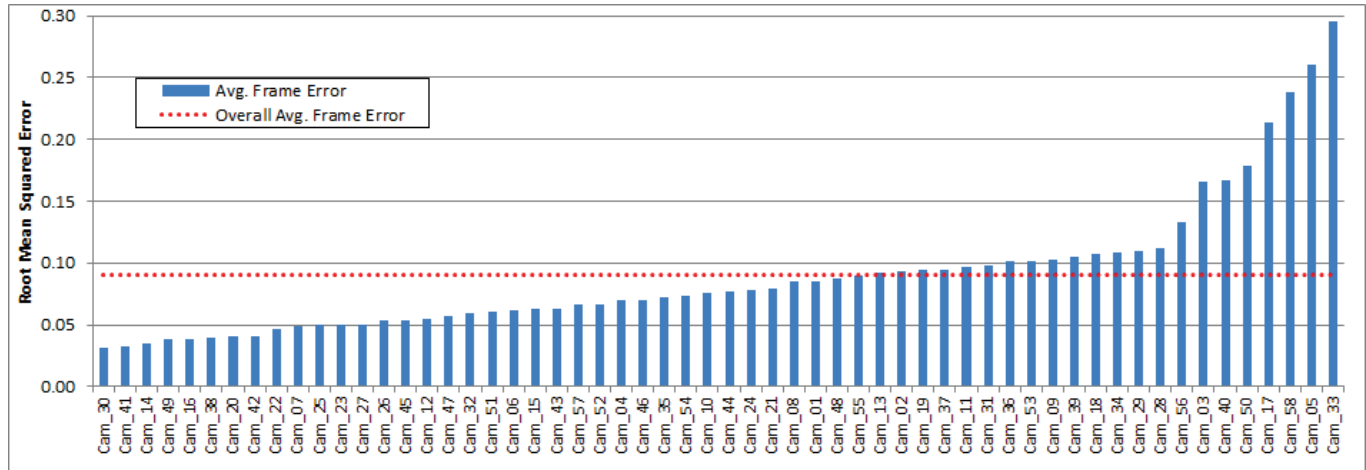


Figure 3. RMS error for all cameras in the laboratory study, using intrinsic parameters from Camera 59.

Table 1. Front camera intrinsic parameters.

Camera	Fx	Fy	Cx	Cy	Distortion
Front	347.2	352.3	241.6	188.9	-0.423287 0.183867 -0.038207 0.000657 0.000187

## Extrinsic Camera Calibration

### Motivation

The extrinsic parameters give a mapping of the image to the “real world,” thereby truly converting the video scene to a measurement device. We sought to determine if traditional methods for extrinsic calibration could be utilized, which largely involve mapping known real-world coordinates of feature points to the image and computing the transformations (translation and rotation matrices) that optimally project the data points. We needed a process, however, that does not require manual location of fiducial points for every vehicle in the study and could deal with the generally poor quality of the video stream. We also sought to determine if any method of calibration (manual or automated) would be effective in terms of accuracy, especially given the variation in the intrinsic parameters, which we are trying to replace with a single model. We noted that the SHRP2 radar system<sup>14</sup> had an error of roughly 0.25 to 0.38 meters, depending on a variety of factors. We believe the video data can add value even at errors larger than this, but we hoped to achieve comparable values. Finally, we were interested in determining if camera extrinsic parameters changed

during the course of the study and the degree to which they changed. Some methods described in the literature<sup>15,16</sup> were not applicable because of various restrictions or the fact that we needed to use data “in the wild.”

### Procedure

For the extrinsic camera calibration, we must map the real-world coordinates of particular fiducial points in a given video frame to their corresponding image points, undistort the image using the intrinsic parameters, and then compute the mapping (using a translation and rotation) from the 3D real-world coordinates to the 2D image plane. While there are ways to possibly automate this process, we investigated other methods to perform this task that we hoped would be less labor intensive.

We started by taking a sample video frame from a trial video segment,<sup>17</sup> used Google Earth<sup>18</sup> to measure the real-world coordinates of roadside fiducials (i.e., lane markings) relative to the vehicle view, and estimated the translation and rotation matrices needed to project the scene to the image. We identified these particular parameters as the “baseline” and surmised that we could compute subsequent parameters based on small changes to these settings. In particular, we determined that the largest likely variation in camera positioning is the “tilt,” the angle relative to the horizon. This is because the DAS unit has a screw that allows this angle to be changed easily for a particular vehicle geometry. Therefore, our first goal was to find a method to correct for the tilt of the camera. We noted that lane markings on the road are common, and therefore we hoped to leverage this fact for our calibration method. The lane markings are generally parallel in the plane of the road, and when projected to the camera sensor they converge. Our goal was to

convert the perspective view to a plane view by finding a homography that could transform lines that are converging in the camera view into parallel lines in the plane of the road.

To this end, we implemented a procedure to find the lane markings and, assuming they comprise a straight line in the plane of the road, estimated the tilt of the camera by finding the angle that translates the markings into a pair of parallel lines. We applied a directed Hough transform<sup>19</sup> to find lines at angles of 30 to 60 degrees to a fixed region 280 by 50 pixels just in front of the vehicle (Figure 4). (We note that the use of Hough transforms to locate lane markings is fairly widespread in computer vision applications for transportation; see, for example, References 20–24.) While lines were not detected in every frame, we found that multiple frames could be used to provide a robust measure of the right and left side lane markings and an estimate of the linear path they made on either side of the vehicle. The tilt angle was then found to be the angle relative to the baseline parameters that caused the lines to be parallel. This simple procedure was easy to implement but did require some manual screening to remove spurious lines that were detected (e.g., shadows from power or telephone lines).



**Figure 4.** Lane markings located by directed Hough transform.

We tested this procedure on data from the VTTI Training Set.<sup>25</sup> This freely available data set consists of three drivers in three different vehicles who traveled many of the same the roads in the Raleigh–Durham area of North Carolina which were traversed in the NDS. We found that the assumption of negligible pan—the side-to-side variation in camera positioning—was frequently erroneous. Fortunately, we could easily correct for this angle by ensuring that the projected lines were vertical in the new projected space.

To test the method on a larger set, and perform a longitudinal study, we obtained front camera view trip data from 10 vehicles at the beginning and end of the study from VTTI. We specified that we desired trips along roadways with lane markings, with north-south or south-north travel near the middle of the day (to reduce solar

glare). Using this data, we computed the pan-tilt corrections to test our method. We also conducted a manual point selection calibration to serve as a comparison by selecting key feature points in an image frame of interest and then using Google Earth to measure the relative coordinates of the points. Care was taken to find points that were consistent from Google Earth and the study video. In at least a few cases, the line markers were in different places due to the time difference between the study video and the Google Earth imagery. To determine the method's accuracy, a different frame was selected to test and compare the manual calibration with the pan-tilt calibration; for this frame, manual fiducials at known locations were again chosen in the image and projected to the real world with the manual calibration parameters and the pan-tilt parameters, and the resulting RMS error was used to compare the accuracy of the methods.

## Results

Our first concern was the stability of the parameters from the start to the end of the study. We found that out of the 10 vehicles, eight exhibited stable parameters from the beginning to the end of the study. The sample size here is small, but it is encouraging that there is a possibility that parameters estimated from the start of the study will function for all trips. Two vehicles had notably different parameters from the start to the end of the study. The change for vehicle 2 can be observed directly from the image, as shown in Figure 5 (vehicle 10, with virtually no change, is shown for comparison), with notable changes in the reflection of the dash and the hood itself at the bottom of the image. (We used the mirror image of the scene for the left image to help illustrate the changes.) A plot of the pan-tilt change for the subset of vehicles is shown in Figure 6.

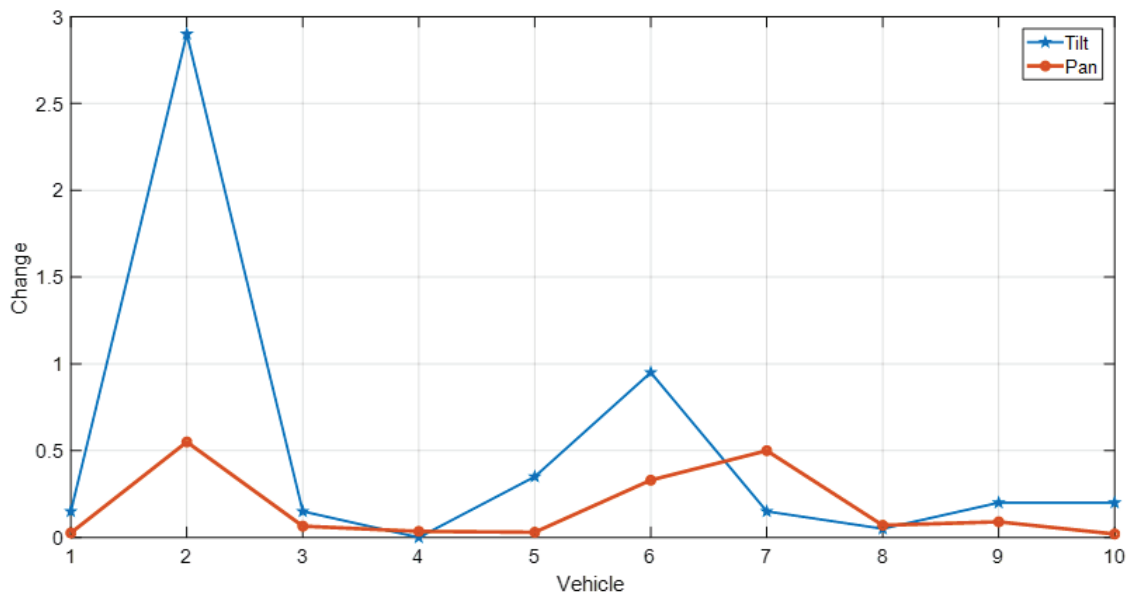
The errors of the two methods are comparable, as shown in Figure 7. Note that there are two measurements per vehicle for the start and end of the NDS study. Overall, as shown in Table 2, the pan-tilt correction gives better results generally. We suspect this may be due to the ability to extrapolate a measurement from the line characteristics, which is more precise than selecting points by hand in the images, which are often rather low quality due to compression, imager quality, and the challenges of real-world imaging outdoors. The error is also comparable to the capabilities of the SHRP2 radar systems.

We also found that on occasion the roll—the rotation of the camera along the axis of the vehicle's direction of travel—was often not negligible and needed to be corrected, at least in some cases. This was apparent from the projected data for one particular case, where the parallel lines on either side of the vehicle had a visible skew after projection. By manually comparing the skew from left to right, we estimated a roll adjustment which resulted in a reduction in the projection error from 0.53 m to 0.24 m. Closer examination of this particular case showed that the roll could be observed in the original data (Figure 8).

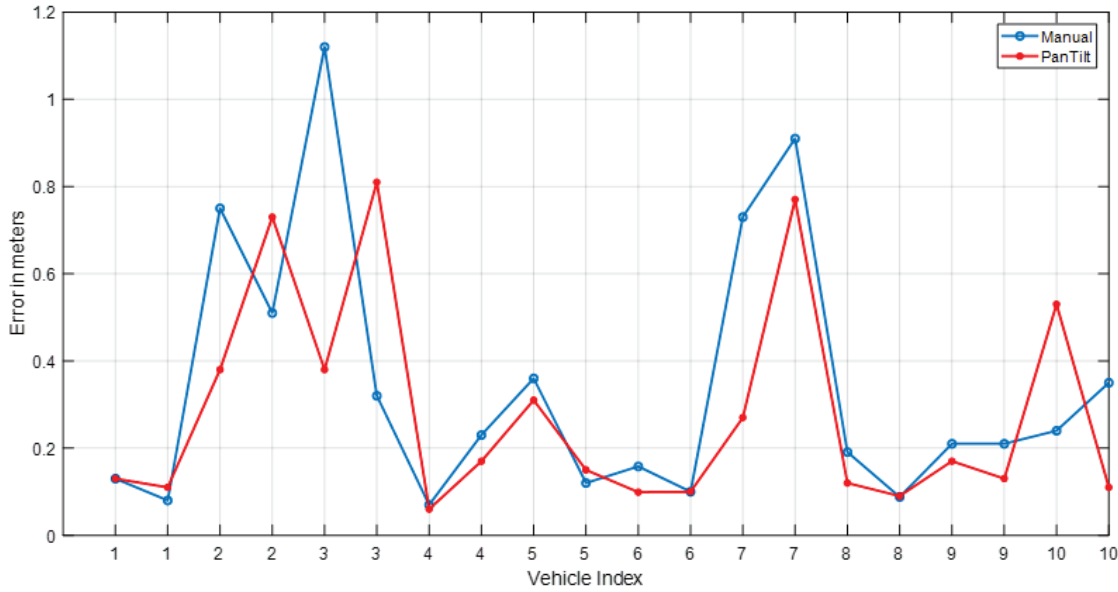




**Figure 5.** Comparison of two vehicles from early and late in the study. The top left shows a frame from early in the study for a vehicle with stable extrinsic parameters; the top right shows the same vehicle but late in the study. The left image is mirrored to show how well it orients with the same features in the bottom of the right image. In contrast, the bottom two frames show a different vehicle early and late in the study, where the camera has clearly shifted orientation, as shown by the misalignment on the features at the bottom of the frames.



**Figure 6.** Change in tilt and pan for each of the 10 vehicles from early to late in the study.



**Figure 7.** Error for each trip in the study, computed by the manual method and the automated pan-tilt method. The manual calibration was computed on a target frame and then applied to a sample frame; the pan-tilt was computed on the same target frame and then applied to the sample frame as well.

**Table 2.** Comparison of errors in the manual and pan-tilt methods.

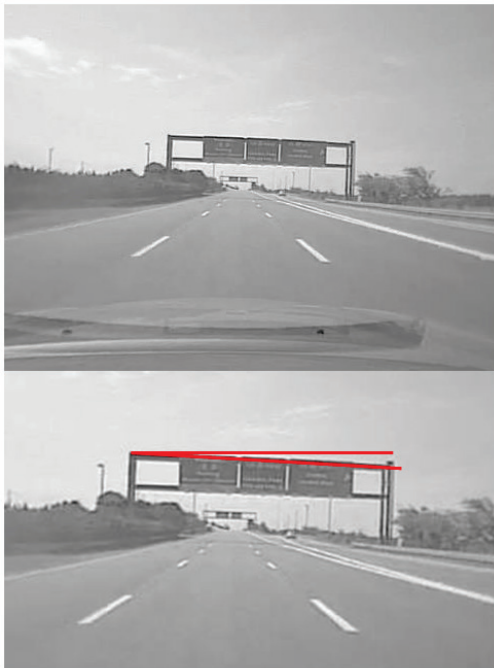
Method	Mean Error	Median Error	Max Error
Manual	0.34 m	0.22 m	1.12 m
Pan-Tilt	0.28 m	0.16 m	0.81 m

## Application Example

The utility of the front video in the study is evident, but an example of the utility of making video measurements is helpful in this context, especially given that the SHRP2 data contains already processed radar data. The concept of combining radar and visual measurements has been explored extensively.<sup>26</sup> Indeed, the radar data is very useful, as it contains tracking information for spatial and relative velocity kinematics for up to eight vehicles per frame. Radar also seems to have a larger effective range than the video data, as the vehicles become difficult to differentiate optically at greater distances, while the radar can still perform detection and tracking (with some constraints, such as the roadway topology). However, the video data has a wider field of view (60 degrees as opposed to ~36 degrees for radar) and can also help differentiate important characteristics such as vehicle type. In Figure 9, we see the result of a video frame from the VTTI Training Set where a vehicle overtakes the participant in the left lane. The vehicle is detected by optical methods (in this case with the YOLO9000 deep network<sup>27</sup>) when the vehicle is very close to the driver's vehicle, where vehicle interactions are most important. In addition, the detection occurs a full 4 seconds before the vehicle is detected by the radar (bottom frame). Measuring the passing vehicle's dynamics with a calibrated camera would contribute to a deeper understanding of the driver's behavior during this event.

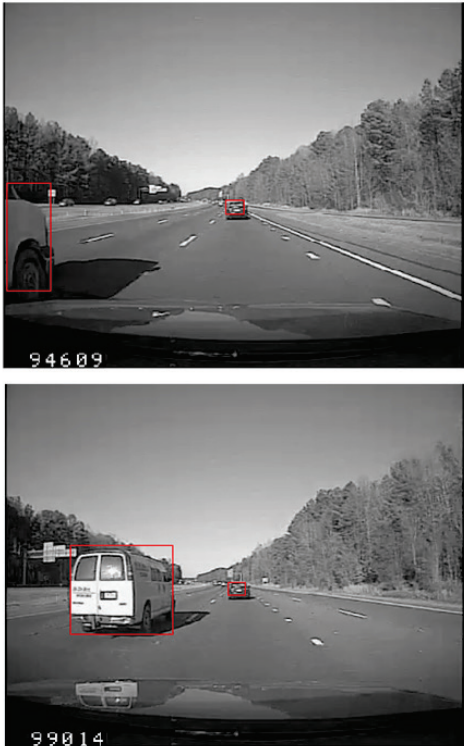
## Conclusions

In this work we described a procedure for estimating intrinsic camera parameters for SHRP2 NDS cameras and applied the parameters to identify a procedure for finding the extrinsic parameters. Correcting for pan-tilt rotation changes using lane markers and automated image processing shows considerable promise, generally yielding values comparable with or better than pure manual calibration. Furthermore, the values are comparable with those obtained using the radar systems in SHRP2 and therefore



**Figure 8.** View from camera with large roll. The overhead sign is perpendicular to the direction of travel but has a noticeable angle with the image frame.

should be of value in use cases of the NDS data where video-based measurements are needed.



**Figure 9.** The vehicle passing on the left (white van) is detected optically by YOLO9000 convolutional neural network in both top and bottom frames but is not detected by the radar until the bottom frame, over 4 seconds later.

There is room for improvement, however, particularly with respect to the automation and processing a large number of trips. Our overall goal is to calibrate the entire SHRP2 dataset of vehicles. While the stability results, we found were promising, with approximately 80% of the vehicles showing stability from the start to the end of the study, our sample set was small, and the number of vehicles with changes was still significant. Thus, we would like to find a way to correct for changes in calibration on the fly, in particular for parameters from the start to the end of the study, such as in these cases. Generally, this work would benefit from some fundamental first-principles analysis of the relative accuracy of point selection and line selection, the latter especially likely, with sub-optimal video data.

We found that pan and tilt are sometimes not enough to characterize a camera. Thus, determining a method to automatically adjust for the roll without using extensive manual measurement is still an area for future investigation. Feature markings that are perpendicular to the direction of travel may offer opportunities for correction, but they need to be preselected and should be sufficiently robust to handle different image conditions and the ravages of time. One approach under consideration relates to travel density maps available through the VTTI Insight website<sup>28</sup>; these maps show roadways where multiple participant vehicles traveled with some frequency. If these roadways are feature rich, they could be used to perform manual measurements on real-world coordinates a single time, and then automated feature selection in the imagery could be used to facilitate the transformation.

Finally, we would like to use the results of this work and further advancements to combine the SHRP2 video data with radar processing. This will require additional alignment and calibration, as it is unlikely that the radar unit was calibrated with the video camera during system installation. An understanding of the variation in measurement error from the video at different ranges would also be useful and will likely include metrics on video detectability.

## Acknowledgments

We would like to acknowledge the support and assistance of Virginia Tech Transportation Institute, especially Jon Hankey, Jon Antin, Suzanne Lee and Miguel Perez.

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>). Work was funded by the Federal Highway Administration of the US Department of Transportation, Exploratory Advanced Research Fund.

## References

- [1] J. Antin, "Design of the in-vehicle driving behavior and crash risk study: in support of the SHRP 2 naturalistic driving study," Transportation Research Board (2011).
- [2] K. Campbell, "The SHRP2 Naturalistic Driving Study", *TR News* **282**, 30–35, (2012).
- [3] InSight Data Access Website, SHRP2 Naturalistic Driving Study, <https://insight.shrp2nds.us/home>
- [4] J. Paone, D. Bolme, R. Ferrell, D. Aykac and T. Karnowski, "Baseline face detection, head pose estimation, and coarse direction detection for facial data in the SHRP2 naturalistic driving study," in *Intelligent Vehicles Symposium (IV)*, 2015 IEEE, pp. 174–179, IEEE (2015).
- [5] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(11), 1330–1334 (2000).
- [6] S. Houben, "Towards the intrinsic self-calibration of a vehicle-mounted omni-directional radially symmetric camera," in *IEEE Intelligent Vehicles Symposium (IV '14)*, pp. 878–883 (2014).
- [7] M. Trivedi, T. Gandhi, and J. McCall, "Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety," *IEEE Transactions on Intelligent Transportation Systems*, **8**(1), 108–120 (2007).
- [8] V. Orekhov, B. Abidi, C. Broaddus, and M. Abidi. "Universal camera calibration with automatic distortion model selection," in *ICIP '07*, pp. 397–400 (2007).
- [9] J. Brito, R. Angst, K. Köser, and M. Pollefeys, "Radial distortion self-calibration," *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 1368–1375 (2013).
- [10] J. Ziegler, P. Bender, M. Schreiber, and H. Lategahn, "Making BERTHA drive - an autonomous journey on a historic route," in

*Proc. IEEE Intelligent Transportation Systems Magazine (ITSM)*, **6**, 8–20 (2014).

- [11] B. Ranft and C. Stiller, "The role of machine vision for intelligent vehicles," *IEEE Transactions on Intelligent Vehicles 1.1*, pp. 8–19 (2016).
- [12] OpenCV website, <https://opencv.org/>
- [13] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV library*, O'Reilly Media, Inc., 2008.
- [14] T. Gorman, L. Stowe, and J. Hankey, "S31: NDS Data Dissemination Activities: Taks 1.6 RADAR Post Processing," March 2015, <https://insight.shrp2nds.us/projectBackground/download/1>
- [15] W. Zhang and M. J. Wolski, "Self calibration of extrinsic camera parameters for a vehicle camera," U.S. Patent No. 8,373,763 (12 Feb. 2013).
- [16] S. Hold et al., "Efficient and robust extrinsic camera calibration procedure for lane departure warning," *Intelligent Vehicles Symposium, 2009 IEEE*, IEEE (2009).
- [17] VTTI Sample video data, available at <https://insight.shrp2nds.us>
- [18] Google Earth, <https://www.google.com/earth/>
- [19] D. Ballard, "Generalizing the Hough transform to detect arbitrary shapes," *Readings in Computer Vision*, pp. 714–725 (1987).
- [20] R. Vivacqua, M. Bertozzi, P. Cerri, F. Martins, and R. Vassallo, "Self-Localization Based on Visual Lane Marking Maps: An Accurate Low-Cost Approach for Autonomous Driving," *IEEE Transactions on Intelligent Transportation Systems*, **19.2**, 582–597 (2018).
- [21] H. Zhou, and H. Wang, "Vision-based lane detection and tracking for driver assistance systems: A survey," in *Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2017 IEEE International Conference on IEEE* (2017).
- [22] B. S. Shin, Z. Xu, and R. Klette, "Visual lane analysis and higher-order tasks: A concise review," *Machine Vision and Applications*, **25(6)**, 1519–1547 (2014).
- [23] B. Fardi and G. Wanielik, "Hough transformation based approach for road border detection in infrared images," *IEEE Intelligent Vehicles Symposium*, pp. 549–554 (2004).
- [24] A. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: a survey," *Machine Vision and Applications*, **25.3**, 727–745 (2014).
- [25] VTTI Training Set sample video data and time series, available at <https://insight.shrp2nds.us>
- [26] P. Banik, *Vision and Radar Fusion for Identification of Vehicles in Traffic*, Dissertation, Virginia Tech, 2015.
- [27] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788 (2016).
- [28] Density map for SHRP2 NDS, <https://insight.shrp2nds.us>

## Author Biography

Jeffrey R. Paone received a BS in Computer Science from the University of Notre Dame (2007), a MEng from University of Colorado (2010), and a PhD from University of Notre Dame (2013). He has worked at Colorado School of Mines since 2015 as a Teaching Associate Professor and Assistant Department Head where his work has focused on computer graphics, mobile applications, and augmented reality.

Thomas P Karnowski received a BS in Electrical Engineering from the University of Tennessee (1988), a MS from North Carolina State University (1990), and a PhD from Tennessee (2010). He has worked at Oak Ridge National Laboratory since 1990 on a variety of research projects in applications of signal and image processing.

Deniz Aykac received a BS in Physics from the Bogazici University, Turkey (1994) and an MS in Biomedical Engineering from The University of Iowa (2000). She has worked at Oak Ridge National Laboratory since 2002 extensively on 3D medical image processing, image and video analysis.

Regina K. Ferrell received a BS in Electrical Engineering (1984) and a MS in Electrical Engineering (1994) from the University of Tennessee, Knoxville. She has worked for Oak Ridge National Laboratory since 1992 on a variety of projects with her primary focus on applications in image processing and big data.

James S. Goddard received a BS in Electrical Engineering from Georgia Tech (1971), a MSEE from the University of Maryland (1974), and a PhD from the University of Tennessee (1997). He has been at Oak Ridge National Laboratory since 1989 specializing in computer vision and image processing research.

Austin Albright received a BS in Electrical Engineering from Tennessee Technological University (2004) and a MS and PhD in Electrical Engineering from the University of Tennessee–Knoxville (2007, 2016). He has worked on a wide variety of research as a member of the research staff at Oak Ridge National Laboratory since 2008. His work as focused primarily on signal processing for, in, and using software-defined radios, on high-speed embedded hardware, and joint time frequency.