

Towards a Practical Stereo Vision Sensor

Mayank Bansal Aastha Jain Theodore Camus Aavek Das
Sarnoff Corporation
Princeton, NJ 08543-5300
{mbansal, ajain, tcamus, adas}@sarnoff.com

Abstract—Development of a practical stereo vision sensor for real-world applications must account for the variability of high-volume production processes and the impact of unknown environmental conditions during its operation. One critical factor of stereo depth estimation performance is the relative alignment of the stereo camera pair. While imperfectly aligned stereo cameras may be rectified in the image domain, there are some errors introduced by both the calibration recovery and image rectification processes. Finally, additional uncalibrated misalignments, for example due to thermal or mechanical deformation in a harsh automotive environment, may occur which will further deteriorate stereo depth estimation. This paper describes an experimental framework for determining these limits using image processing algorithms, operating on graphically synthesized imagery, with performance envelope validation on real stereo image data.

I. INTRODUCTION

Real-time stereo vision sensors have proven to be effective in applications such as off-road robot navigation [1], terrain map building [2], autonomous vehicle detection [3] and 3D scene reconstruction for tele-immersion [4]. Use of such a sensor in an automotive environment, however, requires a quantitative analysis of the robustness of the stereo depth estimation to manufacturing defects and environmental influences (such as thermal deformation) on the stereo rig. In this paper, we consider a parallel camera stereo configuration being used in an automotive setting for target range estimation. The errors introduced by the manufacturing process in the production of an ideal stereo rig necessitate initial calibration of the rig. The camera misalignments introduced in the vibration and thermal deformation prone automotive setting coupled with the calibration recovery errors may lead to large depth recovery errors, affecting the efficiency of a stereo vision sensor. Consequently, it becomes essential to know the misalignment bounds within which the system is assured to produce sufficiently accurate results for depth recovery of objects at typical ranges of operation (0 to 30m in front of the vehicle) for applications like collision avoidance. This paper proposes an experimental framework for conducting an analysis that gives an estimate of these bounds for a stereo vision sensor. The framework uses graphically synthesized imagery to arrive at these bounds and hence gives a practical analysis approach. Further, our use of synthetic imagery is justified by validating them against a set of real images for a particular car model.

II. RELATED WORK

There is a significant amount of published research on characterization of range estimation errors based on system

parameters for stereo vision including [5], [6]. Most of this analysis was carried out from the perspective of theoretical modeling and categorization of errors for the stereo vision based range computation process itself. The evaluation measures used to quantify these errors, however, remain useful for a broader study of miscalibration and misalignment effects as performed in our work. Given the gamut of advanced stereo vision algorithms reported in literature [7], even if we restrict ourselves to the computationally efficient ($> 15Hz$) ones, we are still faced with the task of not just evaluating the performance of each algorithm but analyzing the robustness of a specific algorithm to external calibration and alignment errors. There is not much work in this area in the vision community.

In [8] a theoretical sensitivity analysis was performed for a real-time stereo vision system. However, to understand the true effects of misalignment in yaw, pitch and roll (YPR) angles of the cameras on a stereo rig, an experimental evaluation is needed to better quantify the effects which cannot be predicted theoretically and to back up the theoretical predictions. This poses the challenge of creating experimental data for arbitrary ranges of camera miscalibration and misalignments along with ground truth range data. Most studies (e.g. [9]) comparing accuracy of stereo depth estimation were done using synthetic data as ground truth depth information is hard to obtain for realistic scenes. However, in spite of the increasing power of graphics based 3D rendering, there has been no formal validation of synthetic images against real images of the same object in the context of stereo vision. Also, the available real scene data sets are very restricted in scope - they do not incorporate a wide variety of poses and distances from the cameras as is required in the automotive setting.

In summary, given a choice of real-time stereo algorithm with known (theoretical) performance on an ideal camera rig, there is still a need to develop an analysis framework using synthetic images (validated against identical real images) to quantify the tolerances of the stereo depth estimation to camera miscalibration and misalignment present in a manufactured rig. We propose to address this issue, which is especially important for automotive vision, in the current work.

III. STEREO VISION SENSOR

Stereo Vision Sensor (SVS) robustness analysis can be carried out with respect to errors in stereo rig baseline, left and right camera focal lengths and YPRs (Yaw, Pitch and Roll) arising due to miscalibration and misalignment.

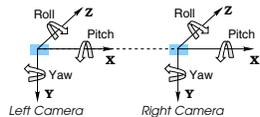


Fig. 1. Stereo Vision Sensor with camera axes and convention for YPR

In this analysis, we focus on rotational errors (YPR) rather than translational (XYZ, which includes baseline) because the former are expected to dominate the latter. However, the framework presented in this paper can be extended to include the other parameters as well.

The SVS used for this analysis is a pair of cameras which are parallel to each other and perpendicular to the baseline (ideal stereo rig). The convention for YPR is as given in the Fig. 1. Our SVS runs *Sarnoff AcadiaTM Stereo* [10] which uses SAD matching windows for computing correspondences.

IV. MEASURES FOR STEREO ANALYSIS

To evaluate the stereo performance under miscalibration and misalignment present in a manufactured rig, we need a quantitative way to estimate the quality of the resulting depth and disparity maps.

The stereo quality metrics can be defined depending on the target application using the stereo output. For example, for an application requiring the range estimate of a target, error in average or median depth is a good metric. However, for a stereo-based vehicle detection and classification application which uses local features for shape estimation, a metric that evaluates error in the local depth variation will be required. In this analysis, we focus on global measures like average and median depth within a predefined Region of Interest (ROI).

For evaluating the effect of miscalibration and misalignment on the stereo performance, we need to compare the stereo output in the misaligned case with that in the ideal case thus precluding the need for ground-truth depth. For this analysis, we define the following stereo quality metrics:

- 1) **Average Depth:** The average of the depth values at the non-void pixels within the ROI.
- 2) **Median Depth:** The median of the depth values at the non-void pixels within the ROI.
- 3) **Standard Deviation of Depth:** The standard deviation of the depth values at the non-void pixels within the ROI.
- 4) **Depth Histogram:** The histogram of the depth values at the non-void pixels within the ROI.
- 5) **Disparity Void Percentage:** The percentage of pixels within the ROI where the stereo returned a void in disparity due to insufficient image texture, low stereo matching energy or a non-unique stereo matching response.

V. THE ANALYSIS FRAMEWORK

In this section, we detail the analysis framework which involves acquisition of real data and generation of synthetic data (section V-A) followed by validation of the synthetic data (section V-B). Then, we determine the theoretical limits of



Fig. 2. Front, Rear, Side, Diagonal-Front and Diagonal-Rear poses of the Dodge Viper: Synthetic (first row) and Real (second row).

acceptable calibration errors and misalignment in section V-C. In sections V-D to V-G, we discuss our experimental analysis methodology for determining acceptable stereo rig *manufacturing* and *environmental misalignment* tolerances. An outline flowchart for the methodology is given in Fig. 11.

A. Data Acquisition and Generation

To estimate the robustness of a stereo vision sensor for practical automotive applications, we need a dataset that corresponds to the actual automotive setting as closely as possible. To arrive at the calibration recovery and misalignment limits, we need a system that allows us to capture data for various left and right camera YPRs without much effort. It is expensive as well as impractical to capture such a huge dataset in the real world, and we thus need to explore a synthetic means of generating data. We however need to validate the synthetic data with the real data to ensure that the stereo performance is comparable in the two cases.

We chose Dodge® Viper RT/10 for our analysis because similar real (scaled model available from Revell-Monogram) and graphical models for the same were easily available. We capture real stereo imagery and generate synthetic stereo imagery (using *OpenGL*® graphics API) of different poses of a Dodge Viper kept at different distances (5.4m to 25.4m) measured between the Dodge Viper bounding box center and the camera. The minimum distance constraint was due to the hood length of the host car (on which the SVS will be mounted) and the (target) Dodge Viper length. This setting closely resembles the scenario where the system will actually be used - both in terms of the target as well as distances. A set of different poses was chosen due to the variability in edge and depth information associated with each pose. A sample set of images from the real and synthetic dataset (with the Dodge Viper at 10.4m) is shown in Fig. 2. For this analysis, we define the ROI to be the rectangular portion of the image occupied by the Dodge Viper.

B. Validation of Synthetic Data

Validation of synthetic data against real data for stereo performance evaluation is essential to ensure that the limits computed by evaluating stereo performance on synthetic data are close to the actual limits.

Synthetic data comprising of stereo views of different poses of the Dodge Viper at different distances, corresponding to an ideal stereo rig, is generated. Real data collection involves emulation of an ideal stereo rig (which is difficult to construct and verify) using a single camera. A single camera is aligned with the world coordinate system and the left image is

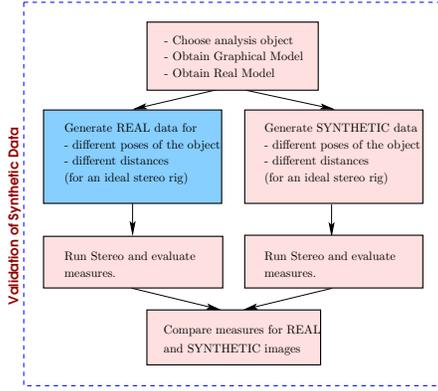


Fig. 3. Validation of Synthetic Data

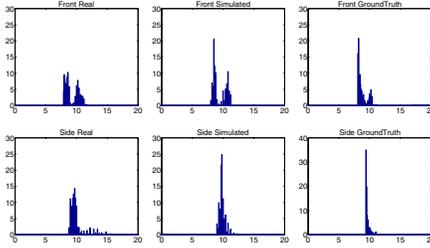


Fig. 4. Comparison of Depth Histograms within the ROI for Real and Synthetic images (Front and Side poses of the Dodge Viper at 10.4m) with Ground Truth Depth Histograms.

captured. It is then translated along the baseline to capture the right image. The stereo procedure is run on the real and the synthetic data and the quality metrics are evaluated on the generated depth and disparity maps. The metrics for the two cases are then compared for validating the synthetic data.

Fig. 3 outlines the basic steps in the validation of synthetic data. Fig. 4 compares the depth histograms for stereo output on the real and synthetic data with the ground-truth depth histograms. The ground-truth was obtained from the synthetic data. The histogram gives a general picture of the depth-variation and the comparison shows that the synthetic data is close to the real data.

Fig. 5 compares the average depth recovered by stereo for real and synthetic images with the ground-truth average depth. This comparison further validates the synthetic data for stereo performance analysis.

C. Theoretical Limits Analysis

In this section, we determine theoretical limits on calibration errors and stereo rig misalignments for a particular threshold on the depth error. This helps constrain the search space for manufacturing and misalignment tolerances. Error in Yaw and Roll of the right camera relative to the left camera introduces an error in disparity due to a horizontal shift (Δx) of the corresponding point in the right image. Error in Pitch and Roll of the right camera relative to the left camera also shifts the corresponding point vertically (Δy) in the right image. This may lead to mismatches or voids in the disparity map during correspondence.

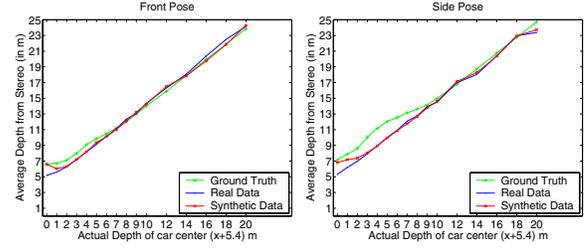


Fig. 5. Comparison of the Average Depth within the ROI recovered by the Stereo for Real and Synthetic images (Front and Side poses) of the Dodge Viper with Ground Truth Depth.

The focal length of the left camera along the x-direction is denoted by f_x (in pixels) and along the y-direction by f_y (in pixels), the stereo rig baseline by b (in meters), the image width by w (in pixels), the image height by h (in pixels), and the depth of the point under consideration by Z (in meters).

- 1) For a **Yaw error** of ΔY (in radians), the depth error (in m) is:

$$\Delta Z = \frac{\Delta Y}{b} Z^2 \quad (1)$$

- 2) For a **Pitch error** of ΔP (in radians), the magnitude of vertical shift (in pixels) for a scanline at a vertical distance of y_c pixels from the image center is:

$$\Delta y = \sqrt{f_y^2 + y_c^2} \Delta P \quad (2)$$

- 3) For a **Roll error** of ΔR (in radians), the maximum depth error (at the top and bottom edges of the image) and the maximum vertical shift (at the left and right edges of the image) are:

$$\Delta Z = \frac{hZ^2}{2bf_x} \Delta R, \quad \Delta y = \frac{w}{2} \Delta R \quad (3)$$

In this analysis, we fix the maximum error in depth to be 1 m for a target at 10 m. Using this depth error threshold, the values of ΔY and ΔR were computed. Assuming that a shift of more than one scanline would deteriorate the stereo quality greatly, the values of ΔP and ΔR were computed (by thresholding on Δy). Since the effect of ΔR is significant only at the image edges, we take the maximum of the two values computed above to study the combined effect in the image interior.

For our SVS, with $f_x = 783.0$ pixels, $f_y = 391.5$ pixels, $b = 0.18$ m, $w = 720$ pixels and $h = 240$ pixels, the values were found to be $\Delta Y = 0.1$ degrees, $\Delta P = 0.2$ degrees and $\Delta R = 0.4$ degrees (after rounding) and will be used as the theoretical limits of magnitude of calibration errors (E_c^t) and magnitude of stereo rig misalignments (L_m^t) for further analysis (Fig. 11).

D. Calibration Recovery Limits Analysis

The stereo rig intrinsics may vary slightly for every unit manufactured. Thus, a calibration step is needed to estimate the unknown values of the rig intrinsics. However, the stereo calibration procedure is itself error prone and the amount of errors introduced in the calibration step directly affect the

amount of external misalignment the rig can tolerate. The more the calibration errors, the less is the external misalignment tolerable for a given degradation in stereo quality. It therefore becomes essential to estimate the amount of calibration errors for a given stereo rig with non-zero YPRs (because of manufacturing variability). In our analysis, we estimate the calibration errors for the *Camera Calibration Toolbox for Matlab* (by Jean-Yves Bouguet at Caltech), which is a practical and robust calibration procedure and can be easily automated.

The Calibration Toolbox requires the stereo images of a checkerboard pattern in different 3D orientations and at different distances from the camera which we generate synthetically using *OpenGL*[®]. The YPRs of the left and right cameras are varied through a wide range of values at a fixed discretization considering all possible combinations. The table below gives these ranges and discretization for our experiment.

Parameter	Test Range (degrees)	Test Discretization (degrees)
Yaw	[-3.0, 3.0]	1.0
Pitch	[-3.0, 3.0]	2.0
Roll	[-3.0, 3.0]	3.0

The calibration procedure is applied to the generated sequences and the YPR recovery error (E_c) for each pair of left and right camera YPRs is computed. Using the theoretical limits on the magnitude of calibration errors (E_c^t) computed in section V-C, the YPR calibration recovery limits ($\pm L_c$) are computed by considering those YPR combinations which satisfy (a) $|E_c| < E_c^t$ and (b) the sum of theoretical depth errors due to errors in each of Yaw, Pitch and Roll recovery (by eqns.1-3) is less than 1 m. For our SVS system, the limits obtained are shown in the table below along with the theoretical depth error sum (described above).

Calibration Recovery Limits ($\pm L_c$) (degrees)			Theoretical Depth Error (m)
Yaw	Pitch	Roll	
2.0	1.0	3.0	0.65
3.0	1.0	1.0	0.80

E. Misalignment Tolerances Analysis

Given a perfectly calibrated ideal stereo rig, it is essential to determine the unknown misalignment that the rig can tolerate before the stereo quality degrades beyond a given threshold. In an automotive setting, there are several reasons why this misalignment may happen e.g. vibrations of the rig mounted in an automobile moving on an uneven terrain or the difference in the thermal characteristics of the environment where the rig is used from those of the calibration environment (leading to expansion/contraction and bending of the rig assembly).

The synthetic data (stereo views of different poses of the Dodge Viper at different distances) corresponding to an ideal stereo rig generated for validation of synthetic data is used for misalignment analysis. A set of right camera YPR misalignments is chosen considering every possible combination of Yaw, Pitch and Roll misalignment within the bounds determined by the theoretical analysis ($\pm L_m^t$) (in section V-C).

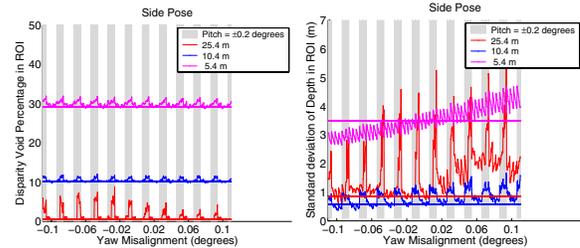


Fig. 6. Disparity Void Percentage and Standard Deviation of Depth within the ROI for Side pose of the Dodge Viper with variation in Yaw, Pitch and Roll misalignments unrolled on the x-axis. For each Yaw misalignment, the Pitch misalignment is varied over the complete range. For each Pitch misalignment, the Roll misalignment is varied over the complete range. Horizontal lines denote the ideal case (without misalignment) disparity void percentage and depth standard deviation within the ROI. The vertical gray bands are the regions where the pitch is maximum.

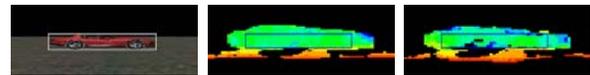


Fig. 7. Left Image, Depth Maps for Ideal case and for a Pitch misalignment of -0.2 degrees ($\Delta Y = \Delta R = 0$) for Side pose of Dodge Viper at 25.4 m from the camera.

Since it is only the misalignment of the right camera relative to the left camera that affects the stereo quality, we pass the generated (ideal) synthetic data through the stereo procedure setting the left camera to be ideal and the right camera to be misaligned within the chosen set of misalignments. We evaluate the quality metrics on the generated disparity and depth maps and determine the misalignment limits ($\pm L_m$) for which the metrics are found to be within the set thresholds.

The table below gives the range of Yaw, Pitch and Roll misalignments along with the discretization at which we performed the analysis. Figs. 6-8 show some of the results obtained from the misalignment analysis.

Parameter	Range (degrees)	Discretization (degrees)
Yaw	[-0.10, 0.10]	0.02
Pitch	[-0.20, 0.20]	0.10
Roll	[-0.40, 0.40]	0.08

Fig. 6 indicates the effect of Pitch misalignment on stereo quality. It can be noticed that the highest peaks in the disparity void percentage and standard deviation depth appear in the gray bands corresponding to the maximum pitch misalignment. This can be attributed to the correspondence errors due to scanline shifts. This effect of change in depth distribution is clearly visible in Fig. 7 where the depth map with Pitch misalignment of -0.2 is compared with the ideal case depth map.

Figs. 8(a) & 8(b) show that the error in average depth varies linearly with Yaw and Roll misalignments, as expected theoretically (eqns.1,3). The variation with Yaw, however, becomes non linear at larger distances (e.g. 25.4m) indicating deviation from theoretical prediction.

Fig. 8(c) gives the combined effect of Yaw and Roll misalignment on the average depth error. It indicates that the line of zero error corresponds to a combination of Yaw and

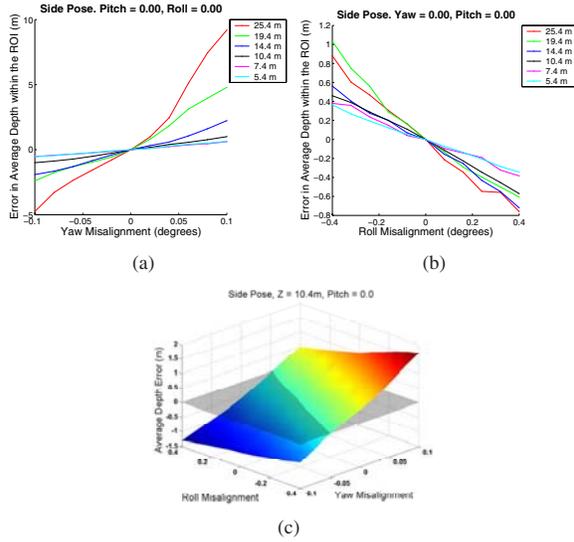


Fig. 8. Variation of error in Average Depth within the ROI (a) with Yaw misalignment ($\Delta P = \Delta R = 0$), (b) with Roll misalignment ($\Delta Y = \Delta P = 0$) and (c) with Yaw and Roll misalignment ($\Delta P = 0$) for the Side pose of the Dodge Viper. (a) & (b) are for different distances and (c) is for the target at 10.4m.

Roll misalignment where they have the same sign. This is because the chosen ROI is in the lower vertical half of the image (not symmetric about the image center). Thus a positive Roll misalignment will increase the disparity while a positive Yaw misalignment will decrease the disparity thus cancelling the two effects and reducing the depth error. Similar is the case when both are negative.

The table below gives the misalignment limits ($\pm L_m$) obtained for our SVS for a depth error threshold of 1 m on average and median depth metrics.

Depth Error Threshold (m)	Misalignment Limits ($\pm L_m$)					
	Using Average Depth (degrees)			Using Median Depth (degrees)		
	Yaw	Pitch	Roll	Yaw	Pitch	Roll
1.00	0.06	0.2	0.08	0.08	0.2	0.08
	0.04	0.2	0.32	0.06	0.2	0.24
	0.02	0.2	0.40	0.04	0.2	0.40
	0.06	0.1	0.16			
	0.04	0.1	0.40			

F. Warping Analysis

The misalignment limits ($\pm L_m$) estimated above are for an ideal stereo rig. In case the rig was found to be non-ideal when calibrated (as is usually the case), there will be an additional loss of stereo quality due to warping of left and right images for alignment. Thus, it is imperative to account for the effect of warping before reporting the misalignment limits. In this analysis, we start by generating synthetic data by considering all possible combinations of left and right camera YPRs (8×8 combinations) at the calibration recovery limits ($\pm L_c$) determined above. The generated data is passed through the stereo procedure setting the left camera at the actual YPR chosen and the right camera at the actual YPR chosen plus a misalignment at the misalignment limit ($\pm L_m$)

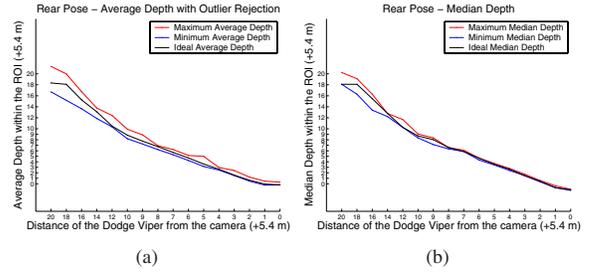


Fig. 9. Comparison of minimum and maximum (a) Average Depth with Outlier Rejection and (b) Median Depth within the ROI for left and right camera YPRs at ± 1.0 degrees with the Ideal case Stereo output for different distances of the Dodge Viper center.



Fig. 10. Left Image, Depth Maps for Ideal case and for a non-ideal stereo rig with Left camera YPR at (1.0, 0.0, 1.0) and Right Camera YPR at (1.0, -1.0, 1.0) and no additional misalignment, for Rear Pose of the Dodge Viper (at 10.4m).

determined above. We evaluate the quality metrics to test if they are within the set thresholds. If yes, then $\pm L_m$ and $\pm L_c$ are the required misalignment and calibration recovery limits respectively. If not, then we repeat the above process after reducing the values for the calibration recovery limits ($\pm L_c$) or the misalignment limits ($\pm L_m$) depending upon where we require looser tolerances.

Figs. 9 and 10 illustrate the effect of warping (without misalignment) on the stereo performance for our SVS system, which uses *Bicubic Warping*, by comparing the quality metrics and the depth maps with the ideal case.

The average depth estimates were found to be quite noisy for a non ideal stereo rig because of a large number of outliers resulting from stereo mismatches. Fig. 9(a) shows the Average Depth estimates after employing a coarse outlier rejection strategy (pruning the points outside a ± 30.0 m window around the actual depth).

Fig. 9(b) shows that the median depth estimate is more robust to warping and is closer to the ideal case estimates. Thus, it is more suitable for practical use - a result which is not evident from a theoretical framework.

An increase in void percentage can be observed by comparing the depth maps in Fig. 10, which correspond to the ideal case and the warping case (with no additional misalignment) respectively. This is because warping leads to texture distortion and aliasing, which may result in low stereo matching energy or a non-unique stereo matching response.

For our SVS system the estimated Calibration Recovery and YPR Misalignment Limits for the median depth metric are shown in the table below. For a stereo rig manufactured within a YPR tolerance given by $\pm L_c$, the misalignment limits ($\pm L_m$) include the calibration errors (E_c) (corresponding to $\pm L_c$) and the unknown environmental misalignments. Thus, the latter is given by $\pm(L_m - E_c)$ and is shown in last column of the table.

Median Depth Error		Calibration Recovery Limits	Misalignment Limits	Calibration Errors	Environmental Misalignment Limits
(m)		($\pm L_c$) (degrees)	($\pm L_m$) (degrees)	($\pm E_c$) (degrees)	$\pm(L_m - E_c)$ (degrees)
1.0	Y	1.0	0.06	0.01	0.05
	P	1.0	0.20	0.02	0.18
	R	1.0	0.08	0.02	0.06

G. Performance Envelope Validation

The misalignment and calibration recovery limits need to be validated for real data to ensure that they were not affected by artifacts in synthetic imagery (e.g. aliasing, single point focus etc). Also, the noise and distortion introduced by the actual camera can throw the stereo quality in either direction.

It is difficult to configure a stereo rig with a given set of left and right camera YPRs (which we want to set at $\pm L_c$) so we emulate a stereo rig using a single Pan-Tilt camera. We configure it for desired left camera YPR (for Roll, we mount the camera on a rotatable axle), take the left image and translate the camera along the baseline direction. We reconfigure the camera for the desired right camera YPR and take the right image. Once the data has been acquired, the remaining steps are the same as the *Warping Analysis*.

For our SVS system, the experiments performed with real data showed that the limits, determined by the warping analysis using synthetic data, were consistent with the real data.

VI. CONCLUSION

By outlining a practical approach to the robustness analysis of a stereo vision sensor, this contribution enables selection of design and manufacturing procedures for a stereo rig with quantifiably predictable performance over time, under varying environmental conditions. The trends observed using this experimental framework prove that a theoretical framework is not enough to gauge the performance of a given stereo vision sensor. Further, it allows identification of stereo output characteristics which are robust to misalignments (e.g. median depth identified for our SVS) and hence can be used by algorithms (which use stereo output) to counter misalignments. To conclude, we hope that this framework will enable the use of stereo vision sensors for automotive safety and convenience.

ACKNOWLEDGMENT

The authors would like to thank Stephen Decker and Salah Hadi from Autoliv Electronics, North America for their contribution and support.

REFERENCES

- [1] A. Reider, B. Southall, G. Salgian, R. Mandelbaum, H. Herman, P. Rander, and T. Stentz, "Stereo perception on an off-road vehicle," in *Intelligent Vehicles Symposium*, vol. 1. IEEE, June 2002, pp. 221–226.
- [2] R. Mandelbaum, L. McDowell, L. Boggoni, B. Reich, and M. Hansen, "Real-time stereo processing, obstacle detection, and terrain estimation from vehicle-mounted stereo cameras," in *Proc. Workshop on Applications of Computer Vision*, vol. 1. IEEE, October 1998, pp. 288–289.
- [3] M. Bertozzi, A. Broggi, A. Fascioli, and S. Nichele, "Stereo vision-based vehicle detection," in *Intelligent Vehicles Symposium*. Dearborn, MI: IEEE, October 2000, pp. 39–44.

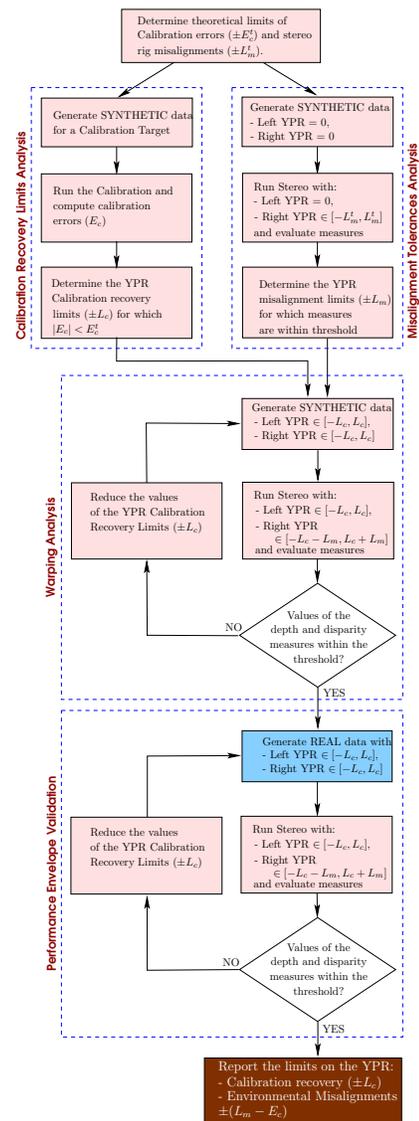


Fig. 11. Flowchart for estimating limits on Camera Misalignments and Calibration Recovery.

- [4] J. Mulligan, V. Isler, and K. Daniilidis, "Performance evaluation of stereo for tele-presence," in *Proc. International Conference on Computer Vision*. IEEE, 2001.
- [5] J. J. Rodriguez and J. K. Aggarwal, "Stochastic analysis of stereo quantization error," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 467–470, 1990.
- [6] S. Das and N. Ahuja, "Performance analysis of stereo, vergence, and focus as depth cues for active vision," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 12, pp. 1213–1219, 1995.
- [7] M. Z. Brown, D. Burschka, and G. Hager, "Advances in computational stereo," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 993–1008, 2003.
- [8] Y. Xiong and L. Matthies, "Error analysis of a real-time stereo system," in *Proc. Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, June 1997, pp. 1087–1093.
- [9] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," Microsoft Research, Tech. Rep. 2001-81, Nov 2001.
- [10] G. van der Wal, M. Hansen, and M. Piacentino, "The acadia vision processor," in *Proc. International Workshop on Computer Architectures for Machine Perception*. IEEE, Sep 2000, pp. 31–40.