

Measuring Natural Scenes SFR of Automotive Fisheye Cameras

Daniel Jakob¹, Eoin Martino Grua¹, Brian Micheal Deegan², Anthony Scanlan¹, Pepijn Van De Ven¹, and Ciarán Eising¹

¹Dept. of Electronic and Computer Engineering, University of Limerick, Castletroy, Co. Limerick V94 T9PX, Ireland

²Dept. of Electrical and Electronic Engineering, University of Galway, Galway H91 TK33, Ireland

Abstract

The Modulation Transfer Function (MTF) is an important image quality metric typically used in the automotive domain. However, despite the fact that optical quality has an impact on the performance of computer vision in vehicle automation, for many public datasets, this metric is unknown. Additionally, wide field-of-view (FOV) cameras have become increasingly popular, particularly for low-speed vehicle automation applications. To investigate image quality in datasets, this paper proposes an adaptation of the Natural Scenes Spatial Frequency Response (NS-SFR) algorithm to suit cameras with a wide field-of-view.

Introduction

Wide field-of-view (FOV) cameras are a unique optical and sensor combination for low-speed vehicle automation [1, 2]. The interest in this sensor modality has spawned several datasets [3, 4, 5], which encourages the development of perception algorithms. However, the optical quality of such datasets is unknown, despite the fact that the connection between optical quality and vision performance has been demonstrated [6].

In [6] an obvious impact of optical lens quality is seen on computer vision performance, where a Cooke Triplet lens degraded the performance of object detection on a BDD100k dataset using a sample of 10,000 images. The BS ISO12233:2017 standard [7] and Slanted Edge algorithm [8] were used to measure optical quality across the spatially variant domain of the camera. In this experiment, test charts were used to measure the optical quality of a camera using three different lens blur models (Gaussian, Superposition, and isoplanar patches). This study was limited to only one lens. This suggests that more information on optical quality is required for publicly available datasets. As of this moment, no datasets provide a means of measuring optical performance using test charts, which suggests replicating optical quality measurements from [6] is difficult and time-consuming. As a strategy to measure image quality performance in the publicly available datasets, the Natural Scenes Spatial Frequency Response (NS-SFR) developed by van Zwanenberg et al. specializes in extracting slanted edges from natural scenes [9, 10, 11, 12]. We adapt NS-SFR to various datasets of spatially variant FOV.

In this paper, we introduce a modified pipeline for generalizing the NS-SFR algorithm to both narrow and wide FOV lenses. The original NS-SFR is unable to eliminate areas not part of the natural scene, which biases optical quality measurements. This especially applies to automotive wide FOV cameras which have ego-vehicle occlusion and lens aperture in the images. Thus, we propose the following contributions in this paper:

1. We propose **regional masking around the scenes** in wide FOV images to remove ego-vehicle occlusion and mechanical vignetting,

2. We propose a **valid and invalid Region of Interest (ROI) Selection System** which chooses edges with the least amount of edge enhancement and noise of slanted edges,
3. We propose using **an Adaptive Radial Distance Analysis** for evaluating optical measurements where radial segments adapted to the regional mask of scene,
4. First analysis across **multiple dynamic automotive scenes** where each scene has a unique camera lens,

Our experiments are focused on four datasets: three are real-life datasets with 90°, 185°, and 190° FOV respectively, and the 4th dataset is based on simulation with 190° FOV.

Context of NS-SFR

For this paper, we use the BS ISO12233:2017 [7] Slanted Edge algorithm as a basis for experiments where the adapted NS-SFR focuses on finding edges on both narrow and wide FOV natural scenes. As far as we know, these experiments are the first attempt at Edge Spatial Frequency Response (e-SFR) for wide FOV cameras without using test charts. More sophisticated cameras exceeding 180° FOV require test charts with greater distortion. There has been little research on taking natural scenes as part of the measurement process for automotive scenes.

Methodology

The proposed Automotive NS-SFR pipeline is depicted in Fig.1. As the goal is to measure the optical performance of a single camera, the dataset must contain images from only one camera. For wide FOV cameras, results are biased where NS-SFR selects the edges that are not part of the natural scene. To avoid result bias, regional masking is used and is applied to the dataset only if parts of the vehicle or camera aperture cover parts of the natural scene, biasing the results. Next, the optimized NS-SFR [9] is applied to the dataset for ROI selection as illustrated by the green box in Fig. 1. The ROI locations of the slanted edge measurements were categorized into three spatial areas on the images: (1) center, (2) middle, and (3) edge. Finally, the mean e-SFR of each radial segment was calculated.

Regional Masks

To generalize processing lenses for NS-SFR, the main interest is to use the natural scene captured by each camera lens. To do this, certain optical effects especially in wide FOV cameras need to be removed which are not part of the natural scene. This involves removing effects such as mechanical vignetting (part of the camera itself is visible in the corners of fisheye images) and ego-vehicle occlusion. These can bias the measurements of NS-SFR where the part of the car to which the camera is attached is visible in any image. All custom masks are created using the Free-

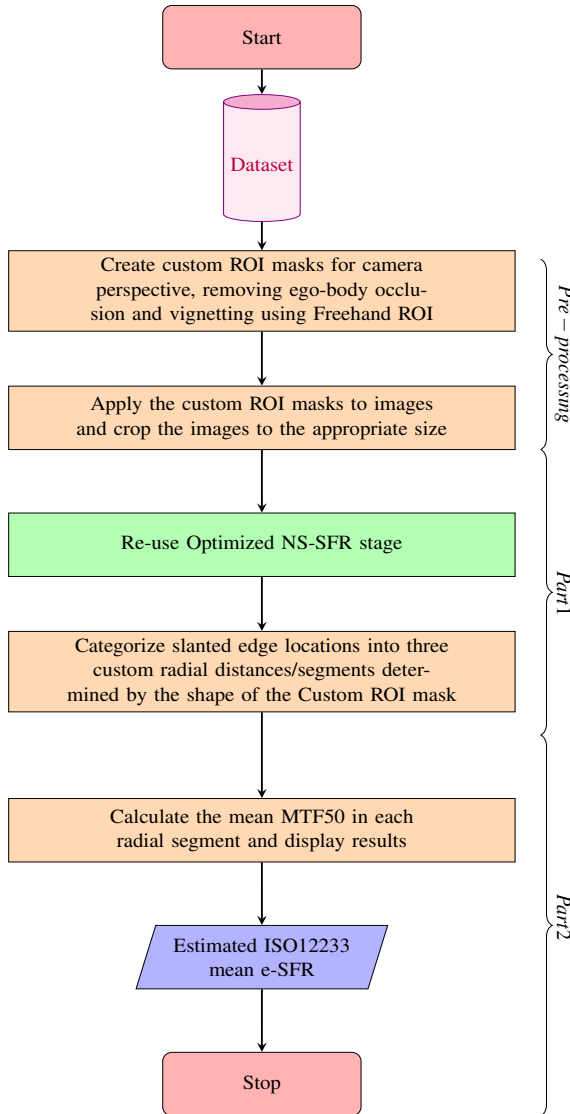


Figure 1: Adapted NS-SFR.

hand ROI tool from MATLAB¹ allowing the user to customize a particular ROI mask to isolate the natural scene. If a new camera is used, a new mask is created to fit the scene(see Fig. 2).

Adaptive Radial Distances

To further analyze the images three adaptive radial distances were created which categorize the location of each slanted edge according to its relative distance from the center of the image. The technique used to create each segment is by taking the Euclidean distance from the central point to the farthest edge of the custom ROI mask. In the case of KITTI, the corner of the image was taken as the farthest distance from the center of the image. The Euclidean distance calculation is represented by the following formula:

$$r_e = \max \left(\sqrt{(x_c - x_m)^2 + (y_c - y_m)^2} \right) \forall (x_m, y_m) \in [X_m, Y_m] \quad (1)$$

¹See Mathworks for information on Freehand ROI: <https://uk.mathworks.com/help/images/ref/drawfreehand.html>

Where, r_e is the Euclidean distance representing the radius of the largest radial distance (i.e. the orange circles (with the exception of KITTI where the orange circle is not visible due to the radius being the distance from corner to centre of the image) in Fig. 2), the point (x_c, y_c) represent the location of the centre of the image or mask, (x_m, y_m) represent points along the periphery of the custom ROI mask $[X_m, Y_m]$ and can be the location of the bottom right corner of the image if no mask is present. (Note: In the case that no ROI is set, $[X_m, Y_m]$ is the periphery of the image). The radii of the radial distances/annuli are obtained by dividing the maximum radius r_e into N proportional parts.

The ratios determined in these calculations are suitable mainly because they give sufficient segments to divide up the images into three categories (i.e., center, middle, and edge) and generalize well regardless of camera type. Ratios smaller than this are unsuitable (i.e., there tends to be bias towards having too many slanted edges favored for one segment).

Re-use Optimized NS-SFR

The ROI selection technique and parameter tuning of NS-SFR optimized algorithm [9] is reused here. Canny Edge Detection with edge masking is used to find edges from the natural scenes and isolate edges of interest [9]. Unlike measurement test charts, NS-SFR does not have a strong sharpening effect on step edges due to the surrounding scene content [9]. The NS-SFR parameters that determine slanted edge selection in natural scenes are the following (the parameters used are given):

1. **Contrast Range** (0.1 - 0.9) - the contrast range between the transition from dark to white in a region with a slanted edge:
 - (a) Low contrast (<0.1) is prone to noise error,
 - (b) High contrast (>0.9) is prone to non-linear sharpening and image processing,
2. **Edge Angle Range** ($0^\circ - 360^\circ$) - there is no restriction on the angle range of slanted edges as it limits the number of edges that can be found in a natural scene except for 0° and 45° which are filtered out and is not possible due to errors [7],
3. **Step Edge Noise Floor (ST)** (0.02 and 0.04 (for images with greater noise)) ensures that the gradient on either side of the edge is uniform. For example, a value of 0.02 signifies a change of pixel value or digital number (DN) of 4.5 for an 8-bit image and 0.04 is twice this value at 9 [11]. It was observed that sharper edges were found as the ST and esfW parameters increased,
4. **Edge Spread Function Width (esfW)** (default value is 5 pixels) remove edges too close together. For example, edges can be 5 pixels apart in an image,

For initial experiments, a few default parameters were used (i.e., ST=0.02, esfW=5 pixels) [11]. These were found to be especially suitable for narrow FOV cameras, such as KITTI. The only difference to the default was removing the Edge Angle range and increasing the default contrast range between (0.55-0.65) to (0.1-0.9) to accommodate for a larger ROI selection. As a second experiment, the default parameters were increased to (ST=0.04,esfW=10 pixels) which showed a higher quality ROI selection for wide FOV cameras (particularly for Woodscape). This was done to understand what happens if distortion is applied to the

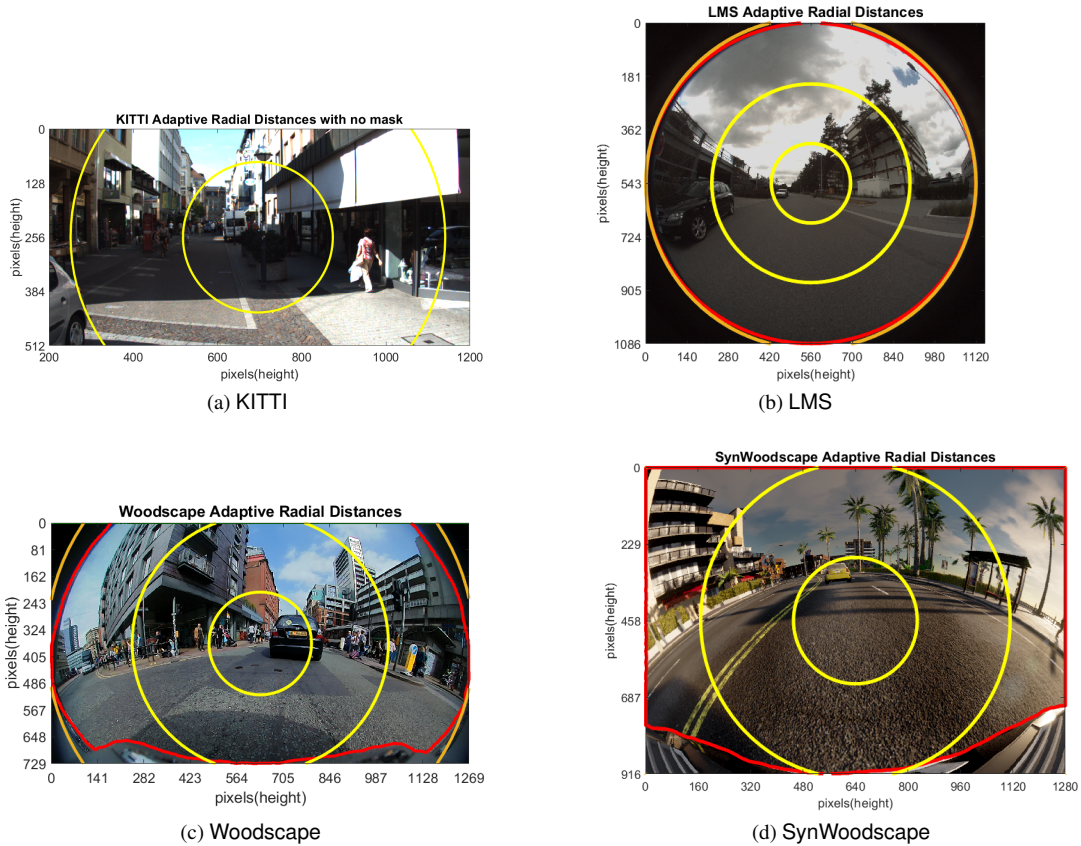


Figure 2: **Adaptive Radial Distances for all four datasets:** (a) KITTI (no mask), (b) LMS (circular mask), (c) Woodscape (Front View mask) and (d) SynWoodscape (Front View mask). Note: all masks are marked by a red outline, inner radial distances (yellow circles) and the outer radial distance (orange circle). For KITTI only the inner Radial Distances are shown between (200-1200) pixels.

images and whether the algorithm performs better with a different parameter combination than the default used in the experiments in the prior art [10]. In the end, we settle on a unique parameter combination for each dataset where each radial segment of the spatial domain is well represented (i.e., each segment should have more than 20 valid slanted edges to calculate the mean).

Valid & Invalid ROI Selection

We apply additional constraints to slanted edge selection where the following constraints were applied:

1. **Local Maxima** < 1.4 SFR - If the local maxima exceed a peak of 1.4 SFR (\cong 25% overshoot), the measurement is discarded. 25% overshoot or less is acceptable² for slanted edges,
2. **Local Minima** < 0.4 SFR - If the local minima exceed 0.4 SFR, the measurement is discarded. Setting this would filter out signals with excessive noise past the Nyquist frequency (0.5cy/px) as shown in Oliver et al. [13],

Measurements that meet the above requirements are marked by a green bounding box and those that do not are marked by a red bounding box as illustrated in Fig. 3a.

²See Imatest website for information on MTF curves and Image appearance: https://www.imatest.com/support/docs/23-1/mtf_appearance/

Radial Distance Analysis

Instead of using three evenly distributed radial distances [9], three adaptive radial distances are proposed dependent on the regional mask. The mean NS-SFR per radial distance is calculated and compared visually between all four datasets. The mean NS-SFR are plotted against each other to gain insight into how the varying spatial domain affects image quality.

Results

Datasets

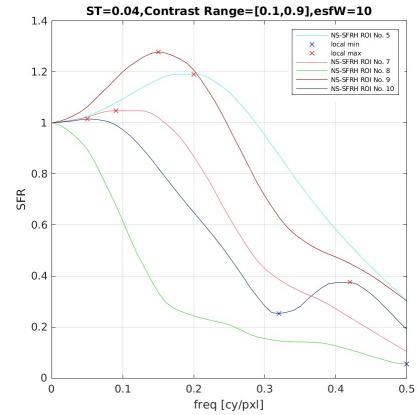
The datasets used in this study are:

1. 1065 Front View 90° FOV KITTI city images [14],
2. 1251 LMS 185° FOV real-life images from Drive A [4],
3. 1514 Front View 190° FOV Woodscape images [3],
4. 500 Front View images of SynWoodscape [5],

Each dataset has one lens calibration and a custom ROI mask applied to the images. The first experiment is to demonstrate the operation of the NS-SFR algorithm on a dataset with little to no distortion (i.e., the KITTI dataset [14]). No mask is applied to these images. The second experiment is to use a circular fish-eye dataset from [4] with wide FOV. A custom circular mask is applied to these images. For the third experiment, a typical automotive dataset, the Woodscape dataset [2] is used in this work for



(a) Valid and Invalid ROI Selection



(b) Measurements

Figure 3: An example of **ROI Selection** from one of the Front View of the Woodscape dataset at ($ST=0.04$, $esfW=10$). Observe that the local maxima (red) and local minima (blue) of each measurement are marked by an x symbol.

wide FOV (i.e., 190° field-of-view) cameras. Finally, SynWoodscape is used which is a synthetic version of Woodscape created in the CARLA simulator using 4th-order polynomial cube-map projection [5, 15].

KITTI camera experiments

ROI selection was iterated over 1065 images of a natural scene in the KITTI dataset and a resulting spatial distribution of the slanted edges can be seen in Fig. 4a. There was no adaptive regional mask required for KITTI. The yellow radial distances from Fig. 2a were calculated by taking one of the corners of the KITTI images which is by default the farthest location away from the centre of each image and the radiuses were determined by a ratio of N proportional segments. The mean MTF of each segment is shown in Fig. 5a.

The LMS Circular fisheye camera experiments

The experiments were repeated on 185° FOV circular fisheye images. A circular mask was created for the dataset illustrated in Fig. 2b where the blue radial distance is calculated based on the furthest point of the circular mask using (1). Fig. 4b represents the resulting spatial distribution of the LMS dataset subsample used in the experiments. In comparison to KITTI, this has a circular aperture and a completely different ROI distribution. The default parameters were used in the KITTI dataset. The mean MTF of each segment is shown in Fig. 5b where all MTF50 measurements are worse than in KITTI (see Fig. 5a). This is the first dataset to exceed the 180° FOV in these experiments.

Woodscape experiments

The Woodscape dataset is composed of 190° FOV cameras [2]. Out of the four camera perspectives available, the Front View perspective was chosen for experiments that consisted of 1514 images. To demonstrate the proposed pipeline on wide FOV, a sample slanted edge ROI selection is shown in Fig. 3 where the green boxes represent valid slanted edge measurements with minimal edge enhancement and noise as specified by the local maxima and minima constraints. As can be seen, NS-SFR has identified 10 potential edges of which only 5 were chosen as valid which

is approximately 50%. Considering the amount of noise natural scenes tend to have it can be concluded that Woodscape has good selection of slanted edges. Slanted Edge measurements of each ROI from Fig. 3a are shown in Fig. 3b where all 5 slanted edges are graphed. The ROI mask of Woodscape is shown in Fig. 2c where the images had both ego-vehicle body occlusion and mechanical vignetting. The experiments resulted in another unique spatial distribution shown in Fig. 4c. For this experiment, both ST and $esfW$ were increased to 0.04 and 10 respectively. The mean MTF of each segment is shown in Fig. 5c.

SynWoodscape experiments

For SynWoodscape experiments, the custom ROI mask can be seen in Fig. 2d. Unlike Woodscape SynWoodscape does not have dark corners from mechanical vignetting and only ego-body occlusion needs to be removed from the images. The spatial distribution of this experiment can be seen in Fig. 4d. By comparing Fig. 4d to Fig. 4c, SynWoodscape has a much denser and wider spread of points than Woodscape using the same parameter settings despite having only 500 images to work with. This comparison proves that in a perfect world simulated scenes tend to have more ideal slanted edges than in real life. For consistency, the same NS-SFR parameters as for Woodscape were used. The mean MTF of each segment is shown in Fig. 5d.

Discussion

In Fig. 4 we show the spatial distribution of slanted edges found by NS-SFR and we can observe distinct patterns emerging where scene dependence affects ROI selection in images with greater distortion. For example, the selection of slanted edges for KITTI Fig. 4a is mostly uniform whereas for wide FOV, ROI selection follows the distortion of the natural scenes. The LMS dataset has a significant number of slanted edges identified on the buildings or in the sky (see Fig. 4b) whereas, for both Woodscape and SynWoodscape, clusters of slanted edges are found along the horizon where the road meets the sky and the buildings (see Fig. 4c and 4d). However, a greater density of slanted edges does not imply better quality selection as is evident in the LMS MTF50 results in Fig. 5b where the quality of slanted edge selection is poor

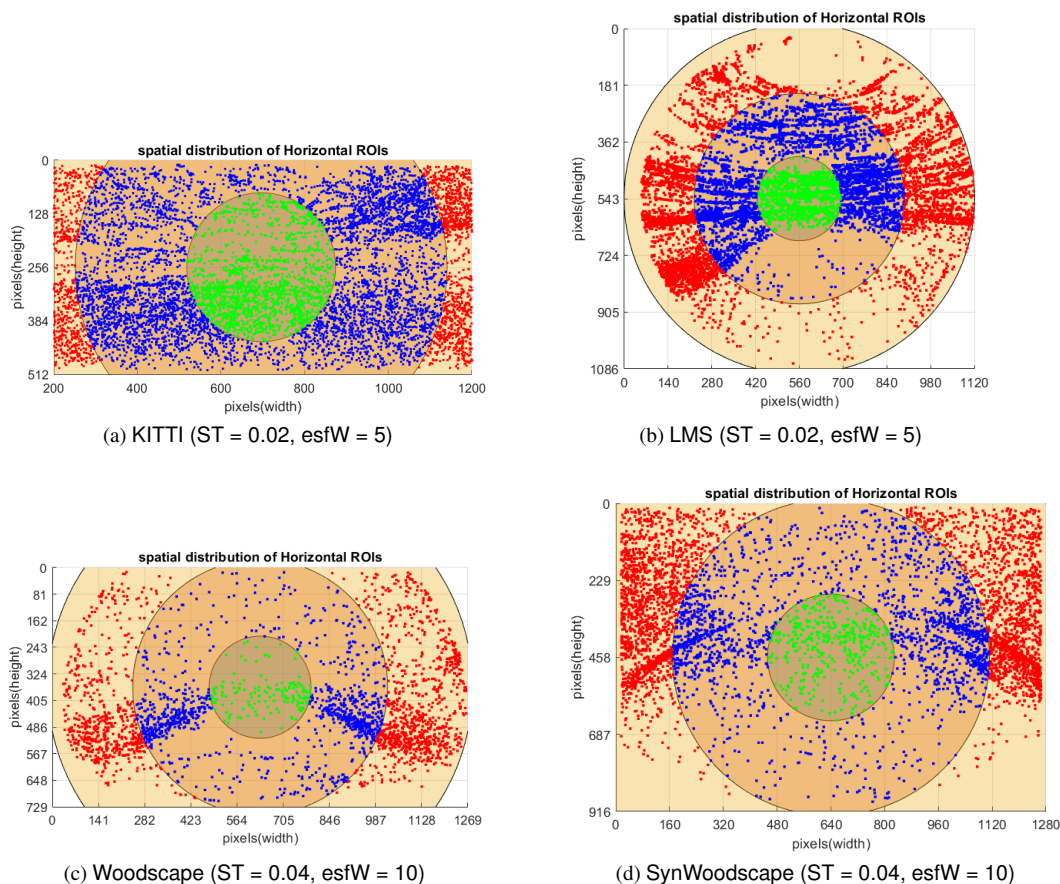


Figure 4: **Horizontal Slanted Edge Locations of all four datasets (refer to Fig. 2 for Adaptive Distances):** (a) KITTI, (b) LMS, (c) Woodscape and (d) SynWoodscape. Distributions show the pattern of ROI selection for NS-SFR for both narrow and wide FOV cameras. Here, we can see the trend of the location of slanted edges in the scenes for the given datasets.

(mean MTF50 is between 0.15-0.18 cy/px). Parameter variations visibly provide higher quality but fewer slanted edge selections for Woodscape (as shown in Fig. 4c) whereas the periphery provides better-slanted edge selection than the centre as shown in Fig. 5c. In contrast to LMS, Woodscape has better MTF50 between (0.31-0.35 cy/px). The SynWoodscape dataset has lower-quality slanted edges than Woodscape suggesting that simulation does not imply higher-quality slanted edges. The central radial segment has the least sharp slanted edges, the middle segment is the second sharpest and finally, the third radial segment has the sharpest edges. Peter Burns et al [16], shows 3rd order and 5th order polynomial measurements on curved edges with similar results. All measurements are of 5th order in these natural scenes.

Future Work

- Consider automated regional masking for each fisheye perspective as originally proposed by Hogan et al. [17].
- Differentiate SFR of different objects in natural scenes.

Conclusion

In this paper, we have demonstrated the adaptation of NS-SFR to automotive datasets. We have demonstrated a novel approach to isolate natural scenes by applying regional masking (where required) and measuring image quality using an Adap-

tive Radial Distance Analysis on four datasets (i.e., KITTI, LMS, Woodscape and SynWoodscape). It was found that Woodscape provided the best selection of slanted edges where MTF50 measurements are the highest out of the four datasets. For horizontal edges, image quality degrades at the centre of the image.

Acknowledgments

This work was supported, in part, by the Science Foundation Ireland grant 13/RC/2094.P2 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero - the Science Foundation Ireland Research Centre for Software (www.lero.ie). The authors of this work would like to thank Oliver van Zwanenberg for assisting in the adaptation of the NS-SFR framework.

References

- [1] V. R. Kumar, C. Eising, C. Witt, and S. Yogamani, "Surround-view fisheye camera perception for automated driving: Overview, survey & challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 3638–3659, 2023.
- [2] C. Eising, J. Horgan, and S. Yogamani, "Near-field perception for low-speed vehicle automation using surround-view fisheye cameras," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 13 976–13 993, 2022.

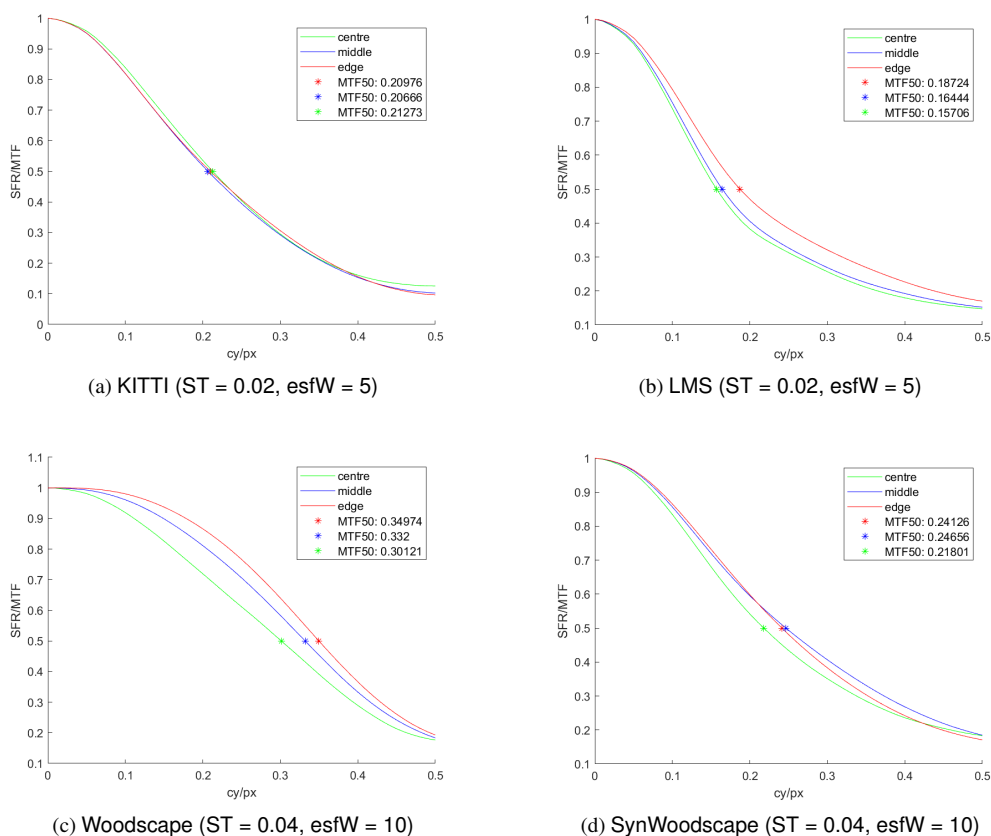


Figure 5: NS-SFR MTF50 Results of four automotive datasets.

- [3] S. Yogamani et al., “Woodscape: A multi-task, multi-camera fish-eye dataset for autonomous driving,” vol. 2019-October, 2 2019, pp. 9307–9317.
- [4] A. Eichenseer and A. Kaup, “A data set providing synthetic and real-world fisheye video sequences,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 1541–1545.
- [5] A. R. Sekkat et al., “Synwoodscape: Synthetic surround-view fish-eye camera dataset for autonomous driving,” *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8502–8509, 2022.
- [6] P. Müller and A. Braun, “Simulating optical properties to access novel metrological parameter ranges and the impact of different model approximations,” in *2022 IEEE International Workshop on Metrology for Automotive*, 2022, pp. 133–138.
- [7] British Standard Institute, *BS ISO 12233:2017 Photography - Electronic still picture imaging - Resolution and spatial frequency responses*. BSI Standards Publication, 2017.
- [8] P. D. Burns and D. Williams, “Camera resolution and distortion: Advanced edge fitting,” *Electronic Imaging*, vol. 30, pp. 1–5, 2018.
- [9] O. Van Zwanenberg, S. Triantaphillidou, and R. Jenkin, “A tool for deriving camera spatial frequency response from natural scenes (ns-sfr),” *Electronic Imaging 2023: Image Quality & System Performance XX*, 2023.
- [10] O. van Zwanenberg, “Camera spatial frequency response derived from pictorial natural scenes,” Ph.D. dissertation, University of Westminster, 2022.
- [11] O. Van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, “Natural scene derived camera edge spatial frequency response for autonomous vision systems,” in *IS&T/IoP London Imaging Meeting*, 2021.
- [12] O. van Zwanenberg, S. Triantaphillidou, R. B. Jenkin, and A. Psarrou, “Estimation of iso12233 edge spatial frequency response from natural scene derived step-edge data,” *Journal of Imaging Science and Technology*, vol. 65, no. 6, pp. 60402–1, 2018.
- [13] O. van Zwanenberg, S. Triantaphillidou, R. Jenkin, and A. Psarrou, “Edge detection techniques for quantifying spatial imaging system performance and image quality,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019.
- [14] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The KITTI dataset,” *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.
- [15] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “CARLA: An open urban driving simulator,” in *Conference on robot learning*, 2017, pp. 1–16.
- [16] P. D. Burns, D. Williams, J. Griffith, H. Hall, and S. Cahall, “Application of iso standard methods to optical design for image capture,” *Electronic Imaging*, vol. 2020, no. 9, pp. 240–1, 2020.
- [17] C. Hogan and G. Sistu, “Automatic vehicle ego body extraction for reducing false detections in automated driving applications,” in *Irish Conference on Artificial Intelligence and Cognitive Science*. Springer, 2022, pp. 264–275.