

# Video-based raindrop detection for improved image registration

Martin Roser

Department of Measurement and Control  
Karlsruhe Institute of Technology

[martin.rosler@kit.edu](mailto:martin.rosler@kit.edu)

Andreas Geiger

Department of Measurement and Control  
Karlsruhe Institute of Technology

[geiger@kit.edu](mailto:geiger@kit.edu)

## Abstract

*In this paper we present a novel approach to improved image registration in rainy weather situations. To this end, we perform monocular raindrop detection in single images based on a photometric raindrop model. Our method is capable of detecting raindrops precisely, even in front of complex backgrounds. The effectiveness is demonstrated by a significant increase in image registration accuracy which also allows for successful image restoration. Experiments on video sequences taken from within a moving vehicle prove the applicability to real-world scenarios.*

## 1. Introduction

Many state-of-the-art computer vision algorithms are designed to work in well-posed visibility conditions. Though this assumption may hold for most indoor and simulation scenarios, environmental noise can badly influence their performance in outdoor settings, e.g. in surveillance applications or in driver assistance systems where a camera is mounted behind the windshield of a moving vehicle.

However, proper operation in the presence of rain is a security-relevant prerequisite to many applications, particularly on board mobile vehicles. For example image registration accuracy declines in the presence of raindrops on the windshield due to mismatched features. In this paper we propose a novel approach to video-based raindrop detection with an application to monocular image registration improvement. Detecting raindrops using a photometric raindrop model also allows for restoring occluded regions by fusing intensity information from nearby image frames as illustrated in Fig. 1.

Traditional image mosaicing approaches [2] fail in registering images from in-vehicle cameras correctly, since they either assume rotations around the yaw axis merely or insignificant scene depth. Our image sequence, however, is recorded at very small frame rates ( $10\text{ Hz}$ ) from within a vehicle moving at  $15 - 25 \frac{\text{m}}{\text{s}}$  which changes scene depth significantly. Furthermore, feature localization noise is high

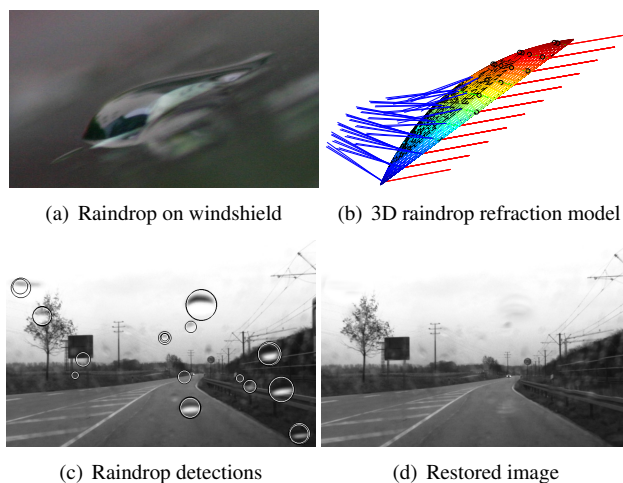


Figure 1. **Raindrop model and image restoration.** (a) shows a typical raindrop's appearance on the windshield. Modeling the refraction of light rays (b) allows for precisely detecting raindrops (c). After registering consecutive frames, occluded areas can be restored using the intensity from neighboring image frames (d).

due to the tilted orientation of the camera with respect to the road and the presence of raindrops on the windshield. This makes direct estimation of the fundamental matrix between consecutive frames difficult [10].

On the other hand, visual SLAM methods [3] are hard to employ due to road homogeneities, pattern recurrences (e.g. lane markings or guardrails) and a wide baseline between frames. Since we are using only monocular sequences, stereo information [15] is not available to us.

For registration we therefore extended the approach proposed in [8], which uses prior knowledge of the vehicle dynamics and camera setup in combination with homography constraints. However, this only works reliably, if enough feature correspondences on the road plane can be matched between consecutive frames. Due to mismatches and occluded image regions this is *not* the case in the presence of raindrops in front of the camera. In this paper we show that automatic raindrop detection helps alleviating this problem.

However, raindrop detection is a challenging problem

due to several reasons: Raindrops on windshields exhibit a large variety in shape and size. Their short distance to the camera results in out-of-focus blur, hereby decreasing distinctiveness of raindrop features. Moreover, transparency and light refraction makes the raindrops appearance highly dependent on the image background. Recently, several approaches have been proposed to overcome these problems:

Garg and Nayar [5, 6, 7] proposed an accurate photometric model for stationary spherical raindrops and determined the effect of camera parameters on image disturbance. However, a static observer or high camera exposure times are assumed making the approach not applicable in scenarios where the camera is moving.

Kurihata *et al.* [12] used a machine learning approach with raindrop templates, so called *eigendrops*, to detect raindrops on windshields. Results within the sky area were promising, whereas the proposed method produced a large number of false positives within the non-sky regions of the image where raindrop appearance modeling becomes more challenging.

Halimeh and Roser [9] developed the basics of a photometric raindrop model to predict the appearance of a raindrop on the windshield. However, they only give experimental validation in a well-defined laboratory setting. They presently disregard the fact that raindrops appear blurred since they are outside the scene focus of the camera.

Yamashita *et al.* [16, 19, 18, 17] proposed several approaches for detecting and removing waterdrops from images. They detect noisy regions using difference images [19] and replace contaminated image parts with patterns from a second camera, assuming a distant scene. Yamashita *et al.* further developed a spatio-temporal approach for detecting adherent noise by image sequence analysis [18, 17], supposing a camera with constant and known yaw rate.

In this paper, we present an integrated concept for raindrop detection (Section 2) that contains the generation of artificial raindrop patterns at *regions of interest* (ROIs), ROI initialization, and raindrop verification by intensity-based correlation. Then we omit disturbing image regions to improve registration accuracy. Thereto we use prior knowledge of the vehicle dynamics and camera setup in combination with homography constraints (Section 3). Finally, we make use of the image registration results and the accurately detected raindrop positions for image enhancement by restoring occluded image areas with intensity information from neighboring image frames. In section 4 we compare the different raindrop detection algorithms in terms of *Precision-Recall* curves and demonstrate improved image registration performance. Finally, we conclude this paper with an outlook on future work (Section 5).

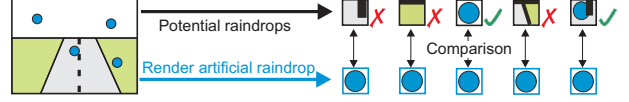


Figure 2. **Flowchart of raindrop detection with RIGSEC.** Artificial raindrop patterns are compared to potential raindrops using intensity-based correlation.

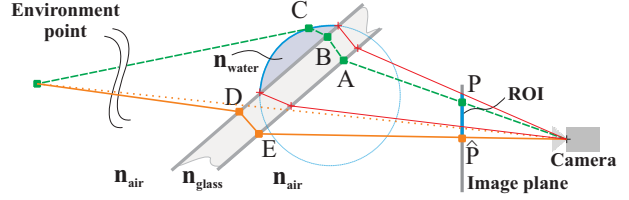


Figure 3. **Refraction model.** Light ray tracing allows for accurate reconstruction of raindrops on transparent surfaces from background scene information.

## 2. Raindrop Detection

Fig. 2 shows an overview over the proposed raindrop detection approach. First, we extract ROIs from the image via interest point detection. In order to check if a region contains a raindrop, we use a photometric raindrop model which renders artificial raindrop patterns for each ROI. These patterns show the raindrop appearance at this location and scale in the image (See Fig. 6). In a verification step, we compare the extracted region with the artificial raindrop pattern.

### 2.1. Photometric raindrop model

In this section, we give a brief overview of the raindrop model we used, called *Raindrop Intelligent Geometric Scanner and Environment Constructor (RIGSEC)* and present all necessary extensions. Our raindrop model is inspired by the model of [9], but we enhance the model by realistic out-of-focus blur and validate the model in real-world scenarios. Furthermore we compare several strategies for finding suitable regions of interest.

Given a potential raindrop, our algorithm models the geometric shape of that droplet on the car windshield. It utilizes its photometric properties and renders an artificial raindrop pattern from points in the environment.

Drop modeling can be performed with varying parametrizations of raising complexity like spherical, ellipsoidal or even more complex models. We use a sphere section, since for most raindrops, it models the shape sufficiently well (see Fig. 1(a)).

Assuming a potential raindrop with radius  $r$  at position  $P = (x, y)$  on the image plane, we determine the part of the scene that will be observed through this raindrop. As depicted in Fig. 3, a light ray emanating from a point in the environment will be refracted by the raindrop and the windshield and finally reaches point  $P$  on the image plane.

The exactly same point can also be seen directly at point  $\hat{P}$ . By tracing the light rays passing through the raindrop, we are able to virtually render the appearance of any raindrop at any location. Snell's law describes the transition of light rays from medium 1 to medium 2. The angle of refraction  $\mu_2$  is a function of the inclination angle  $\mu_1$  and the refractive indices  $n_1$  and  $n_2$ . Hence, for all transitions  $A - E$ , the angle of refraction  $\mu_2$  can be calculated as

$$\mu_2 = \arcsin \left( \frac{n_1 \sin \mu_1}{n_2} \right). \quad (1)$$

Thus, an accurate geometric relationship between  $P$  and  $\hat{P}$  can be derived. Fig. 6(a) shows the sampling of points in the world for rendering an artificial raindrop as determined by our refraction model.

However, in order to compare the intensity values of the mapped environment to the potential raindrop, a prediction of the raindrop photometry is equally essential as the geometric ray tracing discussed above. According to [11], Fresnel's reflectivity coefficients for partly polarized sunlight in the atmosphere going from medium 1 to medium 2 can be expressed as

$$R_{12} = \frac{1}{2} \left( r_{12\parallel}^2 + r_{12\perp}^2 \right) \quad (2)$$

$$r_{12\perp} = \frac{n_1 \cos \mu_1 - n_2 \cos \mu_2}{n_1 \cos \mu_1 + n_2 \cos \mu_2} \quad (3)$$

$$r_{12\parallel} = \frac{n_1 \cos \mu_2 - n_2 \cos \mu_1}{n_1 \cos \mu_2 + n_2 \cos \mu_1} \quad (4)$$

where  $\mu_1$ ,  $\mu_2$ ,  $n_1$ , and  $n_2$  are the angles and refractive indices of media 1 and 2 respectively. Since the pixel intensity  $I_{\hat{P}}$  is known from the camera image, the estimated raindrop intensity at point  $P$  can be determined as

$$I_P = \frac{I_{\hat{P}}}{\prod_i (1 - R_i)^j}. \quad (5)$$

Here  $R_i$  denotes Fresnel's reflectivity coefficients at points  $A - E$  and  $j = \pm 1$  stands for the direction of the intensity prediction. The geometric relations and the photometric properties allows an artificial raindrop pattern generation as depicted in Fig. 6(c).

The sharp optical imaging of an object onto a camera sensor by means of a thin lens can be described by

$$\frac{1}{f} = \frac{1}{b} + \frac{1}{g} \quad (6)$$

where  $f$  is the focal length of the camera,  $g$  is the object distance and  $b$  is the distance between lens and image plane. Since video sensors for environment perception are

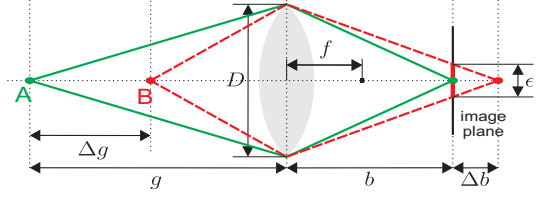


Figure 4. **Optical path for out-of-focus imaging.** Objects that are out-of-focus are imaged blurred. If the camera is focusing point A, point B is imaged onto a disc with diameter  $\epsilon$  in the image plane.

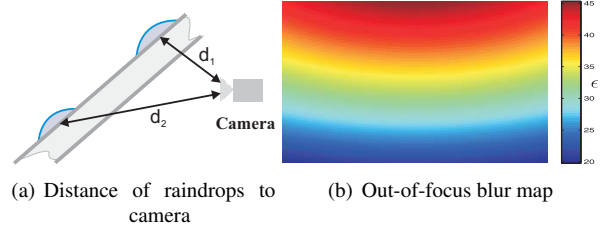


Figure 5. **Computing out-of-focus blur.** The distance of raindrops to the camera is dependent from their position on the windshield (a). This leads to an out-of-focus blur map where each image pixel is mapped to its corresponding unsharp diameter  $\epsilon$  (b).

adjusted for sharp optical imaging of distant objects, close scene points like raindrops on the windshield violate this assumption. They are not imaged sharply, but rather projected onto a disc with diameter  $\epsilon$  in the image plane (see Fig. 4). We model this effect by applying an out-of-focus blur using a disc kernel. According to Fig. 4, the lens equation for out-of-focus objects is given by

$$\frac{1}{f} = \frac{1}{b + \Delta b} + \frac{1}{g - \Delta g} \quad (7)$$

where  $\Delta g$  is known as *depth of field*. Using the geometric relation

$$\frac{\epsilon}{D} = \frac{\Delta b}{b + \Delta b} \quad (8)$$

with lens diameter  $D$ , equation 6 and 7 we get an expression for the disc kernel diameter  $\epsilon$

$$\epsilon = \frac{\Delta g f^2}{O(g - \Delta g)(g - f)} \quad (9)$$

where  $O = \frac{f}{D}$  is the camera aperture size. Since the distance and relative orientation of the windshield with respect to the camera is known,  $\epsilon$  can easily be computed for each point on the windshield as shown in Fig. 5. The blurred raindrop pattern is illustrated in Fig. 6(d). Comparison to the original raindrop (Fig. 6(b)) shows good visual agreement.

Rendering raindrops at multiple scales and all image locations is computationally expensive. On the other hand,

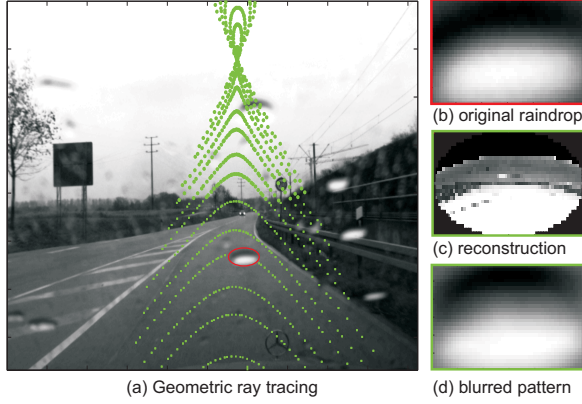


Figure 6. **Artificial raindrop generation.** (a) all pixels of a raindrop at a specified location and scale (red circle) are projected to the environment by our raindrop model (green dots). In (b) the original raindrop is shown, while (c) depicts our reconstruction using all points from (a) and photometric constraints. The result of applying out-of-focus blur to our reconstruction is shown in (d).

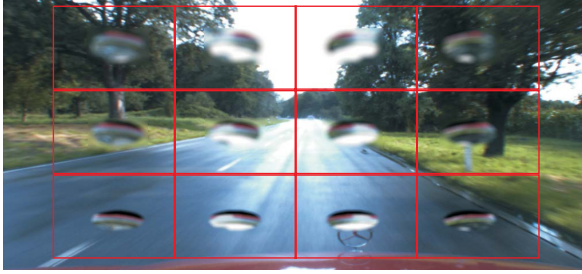


Figure 7. **Fast RIGSEC** algorithm with  $4 \times 3$  artificial raindrops.

raindrop appearance is highly dependent from its background scenario, which lets simple filter masks or *eigen-drops* as discussed in [12] only perform insufficiently. For this reason, we developed a fast RIGSEC implementation that synthesizes a limited number of equally spaced raindrop templates and correlates them in a surrounding region. Fig. 7 shows artificial raindrop patterns and the correlation areas for a  $4 \times 3$  grid.

## 2.2. Initialization

In order to further narrow the search space we first find ROIs using two different interest point detectors described in this section. In this paper we choose the SURF feature detector [1] as a baseline, which approximates the second order image derivatives with box filters. By looking at the scale of the SURF filter response we also get an estimate for the raindrop size in addition to its location.

Gradient based feature detectors like SURF are not performing well for detecting raindrops, because blurred regions only respond weakly. On the other hand, decreasing the detection threshold leads to many false detections that must be rejected by the raindrop model. For this reason, we

present a novel method for detecting blurred image regions by adaptive bandpass filtering. The intuition behind this is that raindrops can be located at specific spatial frequencies which are upper bounded by their out-of-focus blur.

We create a bandpass image  $\mathbf{B}$  for each image  $\mathbf{I}$  by computing the *Difference of Gaussians* (DoG) for each pixel location according to

$$B(x, y) = (\mathbf{I} \otimes g_{\sigma-\Delta\sigma})(x, y) - (\mathbf{I} \otimes g_{\sigma+\Delta\sigma})(x, y)$$

where  $\otimes$  denotes the convolution operator and  $g$  is a 2D-Gaussian. Here  $\sigma$  is taken from the out-of-focus blur map in Fig. 5(b) and  $\Delta\sigma$  has experimentally been determined as  $5px$ . ROIs are extracted by thresholding the bandpass image and finding segments using a connected component algorithm with a 4-neighborhood. Finally, an eigenvalue decomposition of the covariance matrix of all pixel positions belonging to a segment is performed. Convexity ratio (segment area to area of convex hull), dominant orientation (first eigenvector), and aspect ratio (first eigenvalue to second one) are extracted to distinguish raindrops from background features. In our experiments typical raindrops exhibit a convexity ratio of  $> 0.9$ , possess an orientation difference of  $< 15^\circ$  to the horizontal and have an aspect ratio between 1 and 3.

## 2.3. Raindrop verification

Verification of a raindrop candidate is performed by comparing the observed image region with the artificial created raindrop pattern in a small surrounding. We maximize the correlation coefficient of their intensity values  $CC_{\text{intensity}}$  as well as the correlation coefficient of their first derivative (gradient map)  $CC_{\text{gradient}}$ . In order to take different raindrop sizes into account, we slightly vary the scale of the artificial raindrop by a factor of  $1 - 1.5$ .

## 3. Registration and restoration

In this section we show how raindrop detections can be used to improve image registration accuracy. Furthermore, we fuse intensity information from multiple views into one single frame to restore image areas occluded by raindrops.

### 3.1. Geometric scene description

Approximating the road surface by planes enables us to describe perspective mappings via homographies [10].

Fig. 8 illustrates the geometric scene description. We keep the road-to-camera transformation matrix  $\mathbf{T}_{RC}$  and the calibration matrix  $\mathbf{K}$  fixed in all frames and estimate the parameters of the road-to-road mappings  $\mathbf{T}_{RR}(\Theta_i)$ . The 6D-vector  $\Theta_i = (r_x(i), r_y(i), r_z(i), t_x(i), t_y(i), t_z(i))$  represents the rotational and translational parameters between frame  $i$  and frame  $i + 1$ . The extrinsic and intrinsic parameters are determining the transformation from the road



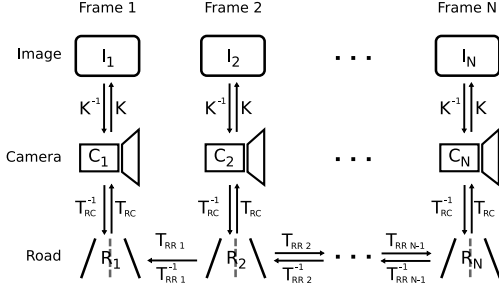


Figure 8. **Geometric scene description.** Transformations used for projecting between different coordinate systems.

to the camera coordinate system  $\mathbf{T}_{RC}$  and the calibration matrix  $\mathbf{K}$ . They are estimated using standard calibration techniques [20] and kept constant over time.

However, only small parts of the scene (e.g. the road) can be represented well by a single plane. We tackle this problem by approximating the world as a box using five homographies, namely the ground plane, the left wall, the right wall, the sky and the background plane. Please note that depth is not represented accurately due to our simplified assumption. However, we model the distance to the left and right wall such that objects of interest (e.g. closeby trees and guardrails) are represented sufficiently well. We achieve satisfactory results in image restoration since only nearby image frames are considered.

### 3.2. Registration

We initially warp the road plane from frame  $i+1$  to frame  $i$  using a prior estimate for  $\Theta_i$  (by means of a first order Markov model) and bilinear interpolation. This projection ensures feature similarity in appearance between frames. We then compute Harris corners in frame  $i+1$  that do not intersect with detected raindrops and search for correspondences in frame  $i$  using normalized cross-correlation. Because of low saliency in our test sequences we compute center-surround feature responses for the maxima in correlation space which are weighted by the template variance. Thereof we only keep the 50% best matches. Further outliers are removed by applying the *Direct Linear Transform (DLT)* algorithm [10] in combination with *RANSAC* [4] which is especially important in situations where other dynamic objects (e.g. vehicles) are present in the scene.

Refining the final parameter set  $\Theta = \{\Theta_i\}_{i=1}^{N-1}$  is done by searching for the MAP solution of

$$P(\Theta|\mathbf{Z}_1, \dots, \mathbf{Z}_N) \propto P(\mathbf{Z}_1, \dots, \mathbf{Z}_N|\Theta)P(\Theta) \quad (10)$$

with respect to  $\Theta$ , where  $N$  denotes the total number of frames. We assume independence of non-consecutive frames and therefore the likelihood  $P(\mathbf{Z}_1, \dots, \mathbf{Z}_N|\Theta)$  factorizes to  $\prod_{i=1}^N P(\mathbf{Z}_i|\Theta_i)$  where the observation probabil-

ity  $P(\mathbf{Z}_i|\Theta_i)$  is assumed to be normally distributed

$$P(\mathbf{Z}_i|\Theta) \propto \exp\left(-\frac{1}{2}\mathbf{d}_i^T \Sigma^{-1} \mathbf{d}_i\right)$$

with covariance matrix  $\Sigma = \text{diag}(\sigma_d^2, \dots, \sigma_d^2)^T$  and  $\mathbf{d}_i$  containing the euclidean reprojection errors between the features connecting frame  $i$  and frame  $i+1$ .

Further assuming independence in the translational and rotational parameters  $(r_x, r_y, r_z, t_x, t_y, t_z)$  over the whole sequence gives

$$P(\Theta) = P(\mathbf{r}_x)P(\mathbf{r}_y)P(\mathbf{r}_z)P(\mathbf{t}_x)P(\mathbf{t}_y)P(\mathbf{t}_z). \quad (11)$$

Here the vector  $\mathbf{r}_x \in \mathbb{R}^N$  contains the pitch angle rate for all frames of the sequence, for example. Ideally one could use a uniform prior for the parameters. However, since in our setting the camera is highly tilted with respect to the registration surface, small registration errors have a large impact on the parameters. Thus we encourage smoothness by putting Gaussian process priors [14] on the function space of the parameters

$$f(i) \sim \mathcal{GP}(\mu_f(i), \sigma_f(i, i')) \quad (12)$$

with  $f(i) \in \{r_x(i), r_y(i), r_z(i), t_x(i), t_y(i), t_z(i)\}$  where  $i$  denotes the frame number and  $r_x(i)$  is the  $i$ 'th entry of  $\mathbf{r}_x$ . For all  $f$  we fix the mean function to an initial estimate of the parameter  $\mu_f(i) \equiv f_0$  and model the covariance function  $\sigma_f(i, i')$  using the squared exponential kernel.

Integrating the prior (11) and the likelihood into (10) and taking the logarithm yields the log-posterior  $\log P(\Theta|\mathbf{Z}_1, \dots, \mathbf{Z}_N)$  which we maximize using standard gradient descent techniques [13].

### 3.3. Multi-band blending

After estimating the parameters  $\Theta$ , reconstruction of image areas covered by raindrops is possible using intensity information from nearby frames. Since simple merging of images leads to unsatisfactory results due to differences in gain, vignetting, object shadows and registration errors we use multiresolution splines [2] which are outlined in the following section.

For each frame one *base image* and one *base mask* are generated. The base image contains all visible pixels warped into the target coordinate system using bilinear interpolation and the homography model from section 3.1, whereas the base mask contains the warped binary raindrop detection images. Combining base images is done by generating *difference masks* and applying multi-band blending [2]. We compensate the gain in each contiguous area independently. This is done by comparing pixels around the area of interest and adjusting the mean intensity value.

Bandpass images  $\mathbf{B}_i^k$  are generated by differencing the previously smoothed base images. Here the index  $i$  denotes

that the image and its mask has been projected from frame  $i$  into the current frame. Please note that only neighboring frames (we typically used a neighborhood size of 6) are considered. For  $k \geq 1$  we have

$$\begin{aligned} \mathbf{W}_i^k &= \mathbf{W}_i^{(k-1)} \otimes g_{\tilde{\sigma}(k)} \\ \mathbf{I}_i^k &= \mathbf{I}_i^{(k-1)} \otimes g_{\tilde{\sigma}(k)} \\ \mathbf{B}_i^k &= \mathbf{I}_i^{(k-1)} - \mathbf{I}_i^k. \end{aligned}$$

where the standard deviation of the Gaussian kernel  $g_{\tilde{\sigma}(k)}$  is set to  $\tilde{\sigma}(k) = \sqrt{2k+1}\sigma$ . Combining overlapping regions

$$\begin{aligned} I_{\Sigma}^k(x, y) &= \frac{\sum_i W_i^k(x, y) I_i^k(x, y)}{\sum_i W_i^k(x, y)} \\ B_{\Sigma}^k(x, y) &= \frac{\sum_i W_i^k(x, y) B_i^k(x, y)}{\sum_i W_i^k(x, y)} \end{aligned}$$

finally yields the mosaic

$$I_{\text{mosaic}} = I_{\Sigma}^{K\sigma} + \sum_{k=1}^K B_{\Sigma}^k$$

with  $K$  the total number of bands.

## 4. Results

In this section we evaluate our approach using two types of experiments: Raindrop detection rates in terms of *Precision-Recall* and image registration errors with and without raindrop detection. The results are generated by applying our algorithm to an outdoor sequence of 302 frames recorded in a rural environment. We used monochrome images with a resolution of  $1024 \times 768$ . The camera system was mounted at a distance of  $\approx 14\text{cm}$  from the windshield with a focal length of 6mm.

### 4.1. Raindrop detection experiments

First, we perform experiments on raindrop detection for the different methods described in section 2. To this end we asked several persons for labeling elliptical regions which appear to contain raindrops within our 302 test frames. Detection results for the different methods are illustrated in figure 9. We present results in terms of *Precision-Recall* for varying detection thresholds. Precision and recall are defined as

$$\text{precision} = \frac{TP}{TP + FP} \quad \text{recall} = \frac{TP}{TP + FN} \quad (13)$$

where  $TP = \text{true positives}$  is the number of correctly discovered raindrops,  $FP = \text{false positives}$  is the number of detections which are no raindrops and  $FN = \text{false negatives}$  is the number of raindrops which are not detected. A detection is considered to be correct if its center is within the true

raindrop. Other closeby detections are removed to avoid multiple true positives for one drop. The ground truth and two exemplary detection results for the image sequence are presented in Fig. 10(a). We further position each frame in *Precision-Recall* space (Fig. 10(b)) and take the mean value for each configuration to generate the *Precision-Recall* plots shown in Fig. 10(c).

Fig. 10(c) compares the performance of the proposed algorithms SURF, BLUR, fastRIGSEC and a combination of BLUR and RIGSEC for varying parameters. SURF is tested with varying filter response thresholds. A recall close to 1 is achieved when keeping the threshold low. However, up to 2000 detected interest points lead to very low precision. Higher thresholds increase precision slowly but at the expense of decreasing recall. In contrast, BLUR shows a non-monotonic behaviour for varying parameters. The curve reaches its best value at a threshold of 0.05. Lower thresholds lead to an insufficient segmentation of raindrops in the binary image and hence show low precision and recall. Higher thresholds omit raindrops that are less distinctive. Recall rates up to 0.67 are achievable with this model. The fastRIGSEC algorithm uses a  $4 \times 3$  grid and we vary the threshold for the intensity correlation coefficient  $CC_{\text{intensity}}$  between 0.1 and 0.95. The best compromise of recall and precision with regards to subsequent image registration was found at values of 0.85 – 0.9. As can be seen, fastRIGSEC dominates SURF and BLUR in performance, verifying the validity of the raindrop model. As depicted in Fig. 10(c), we further combined BLUR with RIGSEC and varied the threshold for  $CC_{\text{intensity}}$ , and  $CC_{\text{gradient}}$  respectively. The presented results reveal that BLUR+RIGSEC clearly outperforms all other proposed methods. However, in terms of recall, fastRIGSEC can exceed BLUR+RIGSEC (at the cost of low precision). This is caused by the fact that RIGSEC only validates interest points. Hence, the recall of BLUR+RIGSEC is bounded by the recall of the BLUR model.

Please also note that neither temporal nor stereo information are used in our system. All tested algorithms can be applied to single monocular grayscale images in contrast to methods using multiple cameras [19] or image sequence analysis [18, 17] for raindrop detection. We believe that further improvements can be made by temporally integrating the respective methods.

### 4.2. Image registration experiments

In this experiment we compare different raindrop detection methods and their impact on image registration. For this purpose we remove all feature detections which are hidden by drops detected with one of the methods described in section 2 and perform image registration as described in section 3.2. We further gather ground truth information visually by maximizing equation 10 using hand selected fea-

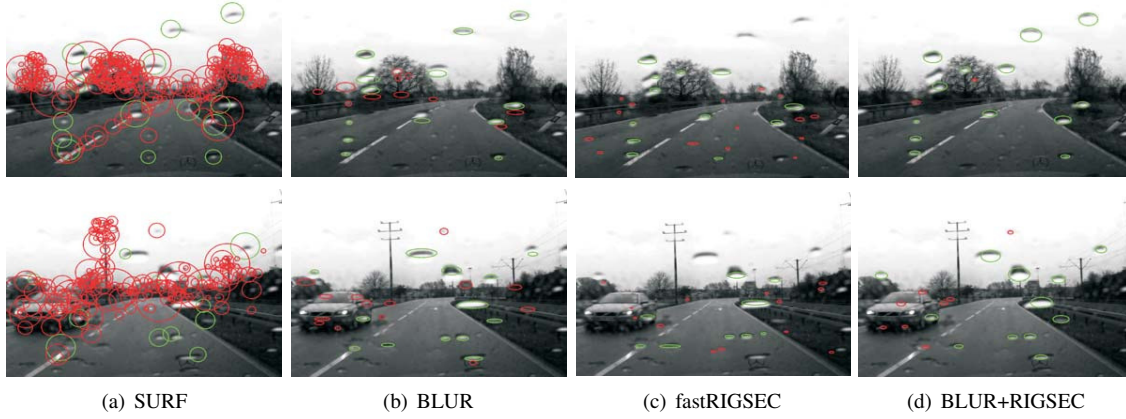


Figure 9. **Qualitative raindrop detection results.** This figure shows our raindrop detection results for the different methods (columns) and two test images (rows). True positives are marked in green, false positives in red.

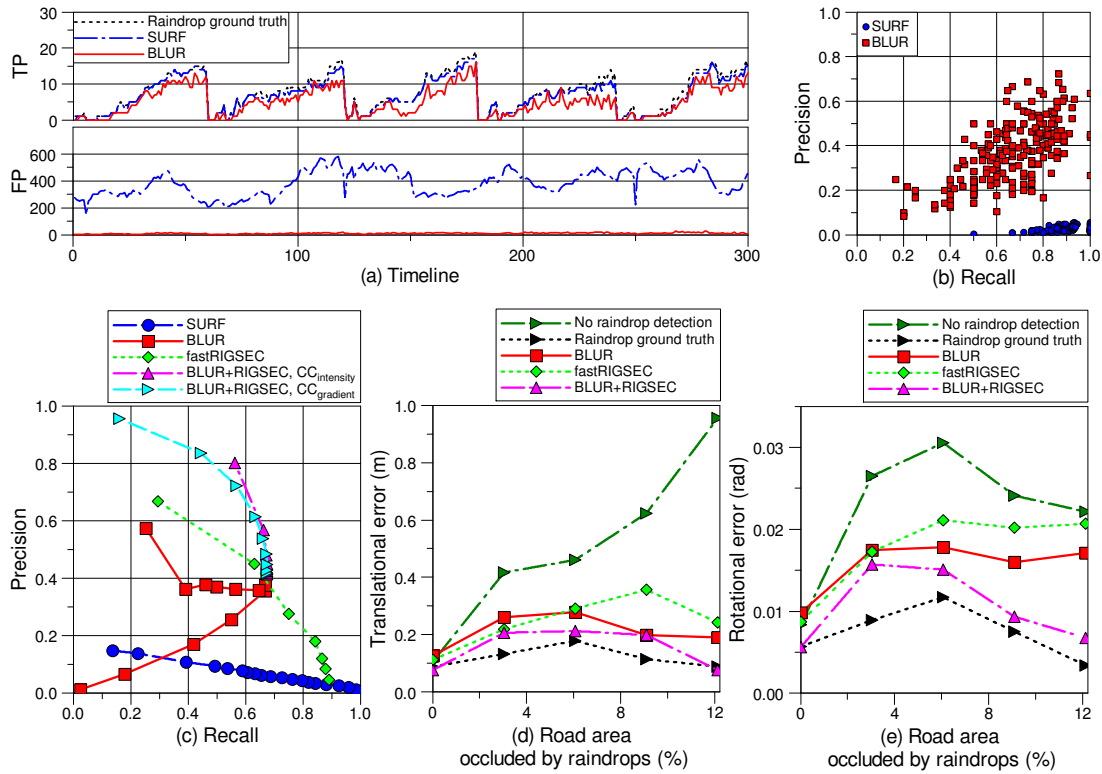


Figure 10. **Quantitative raindrop detection and image registration results.** (a) shows the raindrop ground truth with TP and FP rate of two detectors over time. (b) shows the results of (a) transformed to *Precision-Recall* space to permit a better comparison. (c) illustrates the performance of raindrop detection algorithms. A combination of BLUR+RIGSEC leads to a precision up to 0.8 at good recall rates of  $\sim 0.67$ . (d) and (e) show the translational and rotational registration error over the percentage of occluded road area for the proposed methods. Image registration accuracy can be significantly improved by considering the raindrop detection results.

tures and prior knowledge about the camera setup and the vehicle speed. Results are presented in terms of the mean translational and rotational error between frames.

Fig. 10(d+e) show error histograms over the percentage of occluded road area for the respective methods. Both error measures increase with the number of raindrops when

no raindrop detection is performed. This is due to the fact that raindrops are matched with high reliability in consecutive frames. Please note that location and appearance of raindrops are not reflected by this figure, leading to non-monotonic curve characteristics. However, using our raindrop detection approach lowers the mean registration er-

ror for all test scenarios. We observed that combining the BLUR detector with RIGSEC performed best, which is confirmed by the results in section 4.1. The ground truth acts as an approximate lower bound on the attainable registration error. It is nearly reached using the combined method which shows the effectiveness of our raindrop detection model on real data. We used the successful image registration provided by this method to reconstruct occluded image areas. An example is shown in Fig. 1(d). Please note that some artifacts have been left mainly in the image center. This stems from the fact that regions close to the vanishing point of the image are unobservable since no intensity information from neighboring frames can be used for reconstruction.

## 5. Conclusion and Future Work

In this paper we have presented a novel approach that improves image registration accuracy in rainy weather conditions by considering raindrop detections. We compared different detectors and combined the complementary BLUR detector with the presented raindrop model. This combination revealed superior performance compared to the SURF baseline. We further showed that restoring occluded areas is possible by considering information from neighboring image frames and leads to improved visibility.

Since the sphere section assumption of our model does not cover all raindrop shapes, we plan to extend the raindrop model flexibility. Making use of temporal dependencies in raindrop location and appearance will further improve raindrop detection rates and hence registration results. We believe that this could pave the path for applying many vision algorithms even in rainy weather conditions.

## 6. Acknowledgment

This work is partly supported by Honda R&D Europe. The authors would also like to thank the Karlsruhe School of Optics and Photonics (KSOP) and the Collaborative Research Center (SFB/Tr28) "Cognitive Automobiles" granted by the Deutsche Forschungsgemeinschaft.

## References

- [1] H. Bay, T. Tuytelaars, V. Gool, and L. Surf: Speeded up robust features. In *9th European Conference on Computer Vision*, Graz Austria, May 2006.
- [2] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, 2:217–236, 1983.
- [3] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse. Monoslam: Real-time single camera slam. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(6):1052–1067, 2007.
- [4] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Readings in computer vision: issues, problems, principles, and paradigms*, pages 726–740, 1987.
- [5] K. Garg and S. Nayar. Detection and removal of rain from videos. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume I, pages 528–535, Jun 2004.
- [6] K. Garg and S. Nayar. When does a camera see rain? In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1067–1074, Oct 2005.
- [7] K. Garg and S. Nayar. Vision and rain. *Internatl. Journal of Computer Vision*, 75(1):3–27, 2007.
- [8] A. Geiger. Monocular road masaiing for urban environments. In *Intelligent Vehicle Symposium, 2009 IEEE*, Xi'an, China, 2009.
- [9] J. Halimeh and M. Roser. Raindrop detection on car windshields using geometric-photometric environment construction and intensity-based correlation. In *Intelligent Vehicle Symposium, 2009 IEEE*, Xi'an, China, 2009.
- [10] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [11] C. Klingshirn. *Semiconductor Optics*. Springer, 3. ed. edition, 2007.
- [12] H. Kurihata, T. Takahashi, I. Ide, Y. Mekade, H. M. Y. Tamatsu, and T. Miyahara. Rainy weather recognition from in-vehicle camera images for driver assistance. In *IEEE Intelligent Vehicles Symposium*, pages 205–210, 2005.
- [13] M. F. Moeller. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6:525–533, 1993.
- [14] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.
- [15] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47:7–42, 2002.
- [16] Y. Tanaka, A. Yamashita, T. Kaneko, and K. T. Miura. Removal of adherent waterdrops from images acquired with a stereo camera system. *IEICE - Trans. Inf. Syst.*, E89-D(7):2021–2027, 2006.
- [17] A. Yamashita, I. Fukuchi, T. Kaneko, and K. Miura. Removal of adherent noises from image sequences by spatio-temporal image processing. *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 2386–2391, May 2008.
- [18] A. Yamashita, T. Harada, T. Kaneko, and K. Miura. Removal of adherent noises from images of dynamic scenes by using a pan-tilt camera. *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, 1:437–442 vol.1, Sept.-2 Oct. 2004.
- [19] A. Yamashita, M. Kuramoto, T. Kaneko, and K. Miura. A virtual wiper - restoration of deteriorated images by using multiple cameras. *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, 4:3126–3131 vol.3, Oct. 2003.
- [20] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. *IEEE International Conference on Computer Vision (ICCV)*, 1:666–673, 1999.