https://doi.org/10.2352/ISSN.2470-1173.2018.17.AVM-348

Raindrop detection considering extremal regions and salient features

Vijay C. S, Radhesh Bhat, Vijaya Ragavan
PathPartner Technology Pvt. Ltd., Bengaluru, India

Abstract

Images captured from vehicle mounted cameras endure various uncontrollable adverse conditions such as rain, dust, smudges etc. These result in artifacts which corrupt the captured scene globally or locally. Raindrop is a commonly observed occlusion in images captured from behind a windshield on a rainy day. While reconstruction is a challenging task, detection of these artefacts is also a non-trivial task as there is no well-defined singular model of raindrop or artifact created by it; neither there is a fixed shape, blur and glare level. In this work we address the detection of artefacts caused by raindrops.

We employ a structural verification approach to identify the regions and exploit the uniformity that can be observed in the occlusion region caused due to collection of light rays in the drop by using maximally stable extremal regions (MSER) on the frame to get initial estimate of the regions. False positives are discarded by filtering the estimates based on area, orientation, eccentricity, roundness and convexity. From our results and observation we can conclude that raindrops indeed form extremal regions, however, detection accuracy using MSER is highly susceptible to false positives. Accuracy can be greatly improved using the proposed techniques on the regions based on our observations. A precision-recall analysis is performed to assess the performance of the method.

Keywords *Raindrop, Extremal region, ADAS (Advanced Driver Assistance System), Surface tension*

Background

Natural scene capture through a transparent medium is undertaken in a number of situations and in some occasions for prolonged periods of time. In fact, light accumulation at the camera sensor is implicitly through such a transparent medium viz. the lens. However, most camera systems make a basic assumption that the scene is contaminant free; that the environment of the scene does not affect the capture drastically and the light rays from the scene are unobstructed. This assumption is violated often, particularly in the case of applications such as ADAS and surveillance, where the scene is captured for long duration of time under all kinds of environments. Some examples of such environmental factors are rain, dust, snow, haze, fog etc. Other sources of corruption could be dust/dirt adherent to lens, or water drops or smudges on the lens. Restoring images captured under the influence of such contaminants is a challenging problem and has been studied to some extent in recent years [1] [2] [3] [4] [5]. However, the problem is still open and a general solution is still elusive. Two strategies can be seen in these works, either the artifact is detected first and then restored or a global reconstruction is attempted. Global reconstruction techniques either identify image properties (statistical, frequency or phase related) and then remove effects of drop by filtering techniques or consider water drops as high resolution structured noise and apply denoising techniques using inversion or use convolutional neural networks. Studying water drop properties is challenging as it is highly environment dependent, and learning methods require very large datasets for training (which can include all forms of drops), which is an arduous process. In this work we are interested in the former, two stage process, and specifically we focus on the initial detection stage. Specifically, we focus on a particular form of distortion caused by water drops stuck to the wind shields of a vehicle. We consider such water blobs as extremal regions and hence employ Maximally Stable Extremal Regions (MSER) for identifying potential rain drop candidates followed by several filtering strategies to find water drops in the image.

Environmental factors have impacted capture systems since its inception. Some of these factors have been studied extensively, for example haze, fog etc. Popular dehazing/defogging works are [1] [6] and [7] where focus is to identify the best dark channel, color disparity and maximum contrast. More recently, convolutional neural networks have been applied for dehazing by estimating the atmospheric light in the scene [8]. Another artifact that has been well researched is rain/snow streaks, seminal work in this area was done by Garg and Nayar [2] and [9]. They proposed models for representing streaks and also ways to avoid capture rain itself. Another seminal work by Barnum et.al. analyzed the characteristics of streaks in the frequency domain and used for detection and replacement [10]. Rain streaks have been studies in several other works such as [3] [11] [12] [13]. While these works focus on the handling distortion in scene itself, they cannot address situations where cause of distortion is present on a transparent medium between camera sensor and scene.

A relatively less researched area is of artifacts that are adherent to the lens, such as water drops, dirt and lens dust. Gu *et.al.* in [14] propose image formation models depicting the optical phenomenon to define the effect of small occlusion on the scene. These image formation models are used to design an inverse problem to reconstruct the scene. However, describing an alpha layer cannot explain the refraction and reflection effects that occur in case of rain drops. You et.al.[4] use a two-step detection and removal process for rain drops using long term transition of scene with a static drop. However, these are not suitable for application like ADAS and surveillance because of the static nature of significant parts of the scene. Eigen *et.al.* [5] train a CNN for restoring dust and rain drops in the image. The working of this approach is considered to be similar to denoising with neural networks but with structured noise. While the approach is more general, it is highly data dependent and is limited to the shape and size of artifacts in the training set. In this paper, we consider the detection stage in the

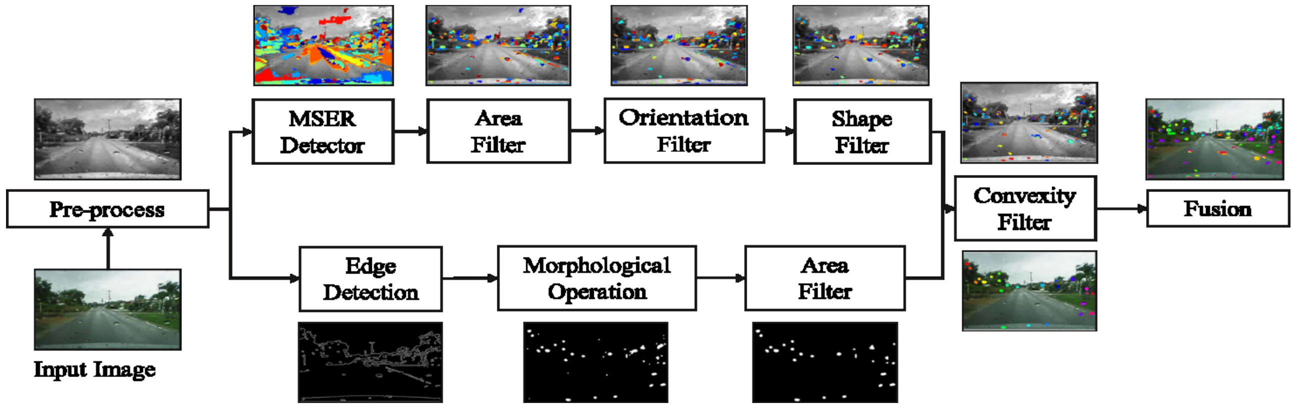

Figure 1. Proposed Method Block Diagram

two step approach which itself is non-trivial and considered a challenge because of the variation in shape, size, color, location and background. In this regard, Einecke et.al. [15] propose a detection algorithm based on normalized cross correlation assuming that the non-occluded parts of the scene are more dynamic. However, this is not true for in-vehicle cameras and surveillance videos as we shall discuss later. Akkala et.al. recently [16] proposed a two-step process for detection where they employ kurtosis for broadly identifying artifact area and then use SVD for detection in road region and DWT for detection in sky region. Although they identify the unique nature of in-vehicle data and handle the road and sky sections separately, the approach restricts performance when detecting many small contaminants and in a fast varying scene.

There are works which have tried to generate photometric models for adherent water drops such as [17]. However, the detection process is computationally intensive and is limited by the type of reference droplet shape that is used. They improved upon this algorithm by using Bazier curves in [18] and were able to overcome shape related limitations, however, the complexity stayed high. A detailed study of state of the art was done by Webser [19] as part of a master's thesis work. The study also proposed a machine learning approach for detecting water drops using texture, shape and context information as a bag of words. However, we would like to contend that more than the texture and context the uniformity within a rain drop can also be exploited to perform better detection. A detection method proposed in [20] exploits this regional nature of rain drop image by initiating the detection process using MSER features [21]. However, we find that their metrics for filtering and reducing false positives is highly limited and that the rationale provided for using MSER is insufficient. We discuss the properties of extremal regions and their applicability in next section and provide extensive steps for limiting false positives. Also another simple procedure is described for identifying small drops which are usually omitted by MSER and hence acts as a complementary to MSER based approach. In a later section we discuss the results obtained in our experiments and later suggest possible extensions.

Proposed Method

In this work we follow a structural verification approach to identify the regions and exploit the uniformity that can be observed in the occlusion region caused due to collection of light rays in the drop. This implies that these form extremal regions i.e. the intensity inside the blobs vary monotonically. The regions are first identified using MSER feature detector. And since MSER detection fails at small regions we employ an edge based detection method which combines edge detection with appropriate morphological operations to identify small contiguous regions. The regions are detected independently and passed through a series of tests to ascertain whether the regions are indeed water drops. These tests are based on size, orientation, eccentricity, elliptical nature, convexity and roundness of the region. The flow of our detection process is shown as a block diagram with corresponding outputs in figure 1. Note that we apply the detection procedure for each frame individually. We did so because we found that multiple frames of a video did not yield significant information for ADAS data. Under ADAS environment, the changes in intensity or motion of raindrops pixels over the frames is not significant. We demonstrate this in the next section.

Extremal regions

The idea of an extremal region was proposed by Matas *et.al.* in [21] and is defined as: for a gray level image I, a region Q, $Q \subset D$ and such that either for all $p \in Q, q \in \partial Q$: $I(p) > I(q)$ (maximum intensity region) or for all $p \in Q, q \in \partial Q$: $I(p) < I(q)$ (minimum intensity region) where p, q are pixels in D set of all connected regions and ∂Q is the boundary of region Q. In other words, any region (set of pixels) in an image is an extremal region if every pixel intensity inside that region boundary is either greater than or lesser than the pixel intensity immediately outside the region. In the case of an adherent rain drops, it is accepted in most of the literature that it will act as a convex lens and focus the scene onto the sensor. Hence, the entire region of drop in image is expected to be more concentrated than the rest of the image. This clearly indicates that these blobs form extremal regions. Also, since we are considering detection at each frame, we consider only those drops which are static and are not in motion. Hence we can make the assumption that the blob will be consistent with a static drop following all physical properties, thus making it maximally stable.

Finding potential regions

Our first step is to identify the potential regions in the image frame that could represent blobs created by water drops. As mentioned earlier, we first attempt to find all these regions using MSER. However, MSER at robust thresholds is unable to detect small regions and many such small regions could be water drops. In order to overcome this limitation, we devise a method to detect small regions based on edges.

MSER based: We pre-process the image frame by adaptively equalizing the image and extract blob features on a gray image. We run maximally stable extremal regions detection on the image frame after adjusting the parameters according to the resolution, field of view and density of rain. Output obtained for couple of images using MSER method is shown in figure 3. Clearly the candidate regions include many non-water drop regions from the scene and also omit small drops in the scene.

Edge based: Detecting small regions is important because in certain situations such as heavy rain, drops form and break up very quickly, or in a light drizzle the drops themselves are small. For this purpose we propose a parallel approach for detecting smaller drops. We observe that drop regions result in strong edges around them due to change in refraction, especially in small droplets. Small drops provide very clear and connected edge on the border of the blob. We exploit this by finding canny edges and retaining only small, fully connected edges. We apply a series of morphological operations to generate dense small regions that could potentially be water drops. We first identify connected components and eliminate large sets. Then perform a morphological filling operation to obtain first estimates of the blobs. We then apply a combination of median filter and image close operation multiple times to remove stray regions. The resulting regions are taken as potential water droplets. Output obtained using edge based method is shown in figure 3.

As expected false positive are very high at this stage in both methods. In next step, we use several metrics to verify observed properties of water drops in the potential regions and select the most probable regions which could represent water drops.

Limiting false positives

Size Rain drops on an inclined surface can be associated with properties such as critical size as larger droplets tend to overcome surface-tension and disintegrate or flow down under gravity as shown in figure 2. The size varies with surface and material. However, a statistical upper bound can be derived empirically.

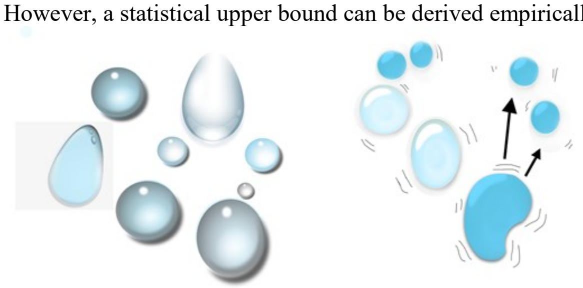

Figure 2. Rain droplets of varying sizes

If the set of regions is $R = \{r_1, r_2 \ldots r_N\}$, N = number of identified regions, then we can say that regions with $Area\ (r_i) > Area_{max}$ $i = 1 \ldots N$ shall be removed from the set. $Area_{max}$ is determined empirically.

Orientation Another observation we make is that the accumulation of water drops cannot happen vertically. Vertical collection reduces the area of contact with surface thus making the drop unstable under the influence of gravity. However this is not true for droplets accumulating horizontally. This property can be applied to the regions detected using MSER which also has an orientation associated with it as an inherent feature. We apply a heuristic upper bound for differentiating horizontal and vertical regions [$Orientation\ (r_i) < \theta$] and select only horizontal regions.

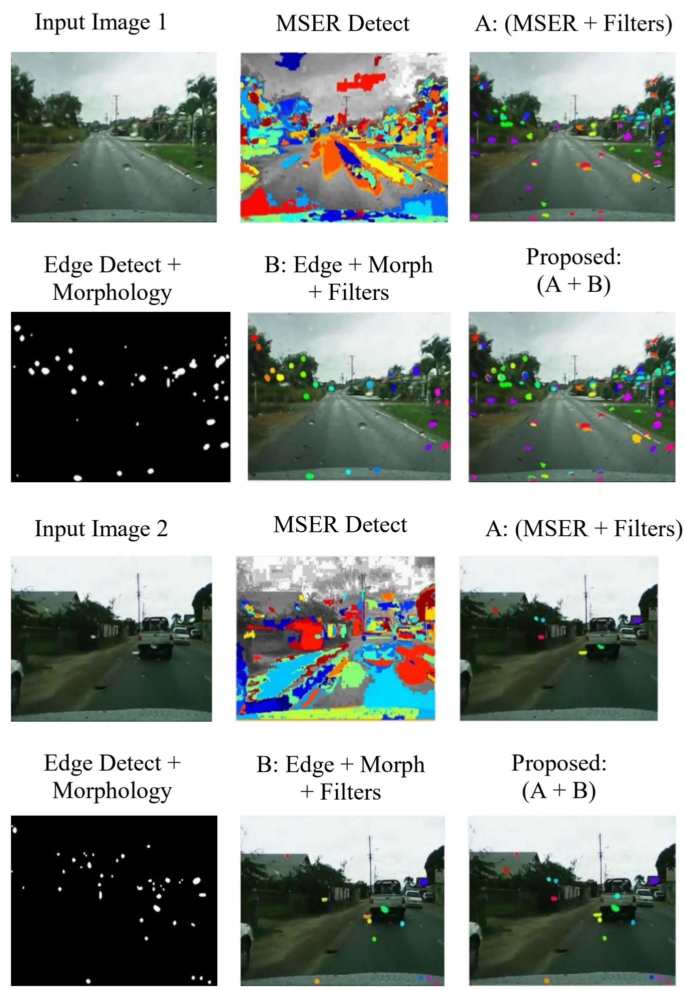

Figure 3. Results of raindrop detection using MSER & Edge based methods, where the colored regions indicate the region detected as raindrops (both true and false detection)

Shape Raindrops spherical structure is due to surface tension of water. Surface tension is the tightness across the surface of water that is caused by the polar molecules pulling on each other. Surface tension is the "skin" of raindrop that makes the molecules stick together. Molecules inside a raindrop are attracted in all directions. Molecules at the surface of the raindrop are being pulled by the molecules next to them and below them. This is depicted in figure 4. We apply three filters to verify the shape of the region. Shape related filters are described below.

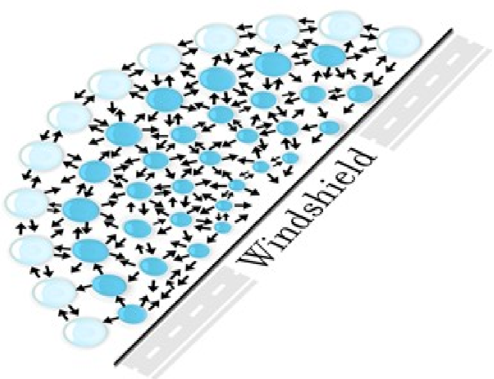

Figure 4. Raindrop surface tension on windshield

Eccentricity There is no restriction on what the shape could be as the largest ellipse that subsumes that region can indicate some features in relation to the region. We calculate the ratio of minor to major axes of such an ellipse and apply a limit to remove thin regions from the set.

Convexity A region is called as convex region, if it contains all the line segments connecting any pair of its point. If the region does not contain all the line segments, it is called concave region. Illustration of convex region is shown in figure 5. This is a property which we consider to be natural to any water drop. The overall surface tension is optimal only for convex shapes and this assumption holds good for windshield. Ideally, to check for convexity one must trace the orientation of tangents around the curve. However we apply a simplistic measure which can indicate convexity. We compare the area of the smallest convex region that contains the detected region with the area of the region. We select those regions which have a measure is less than the threshold which is empirically derived.

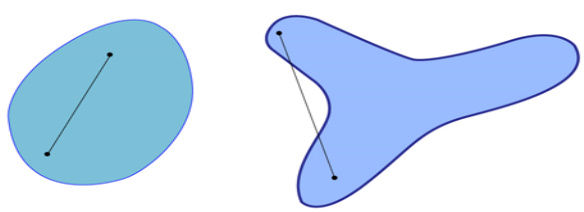

Figure 5. Convex region versus concave region

Roundness An extension of convexity is the round nature of the blob. Although a strict circle is not anticipated but they can be expected to belong to the family of ellipses. This measure indicates the compact nature of the region. We calculate it as

$$\text{Roundness}(r_i) = {Perimeter^2}/{(Area \times 4\pi)}$$

Where $i = 1 \dots N$

A combination of these filters works well for many of the situations. Output obtained is captured in figure 3. More results are shown in the next section.

Experimental Results

This section describes the experiments conducted to measure the accuracy of our proposed rain drop detection method. For this experiment, we validated algorithms with video sequences captured from in-vehicle camera. These input video sequences are taken from public space. We evaluate precision & recall of the raindrop detection methods. Precision is defined as the number of correct detections divided by the total number of detections. Recall is defined as the number of correct detections divide by the number of detectable raindrops.

We experimented with couple of conventional inter frame based raindrop detection methods. First method we took for analysis is based on Time gradient [4]; it assumes that the temporal change of intensity of raindrop pixels is smaller than that of non-raindrop pixels. Second Method we analyzed is based on SIFT flow [4] and it assumes the motion of raindrop pixels is slower than that of non-raindrop pixels. We evaluated these methods with images taken from in-vehicle camera. During our evaluation, temporal data was accumulated over 100 frames. The result is shown in figure 6. We observed high number of missed detections and false detections. The assumptions “motion change and temporal intensity change” of raindrop pixels is slower than that of non-raindrop pixels holds well in constrained environment. But in real road environment the difference in “motion change and temporal intensity change” between raindrop versus non-raindrop pixels is very minimal.

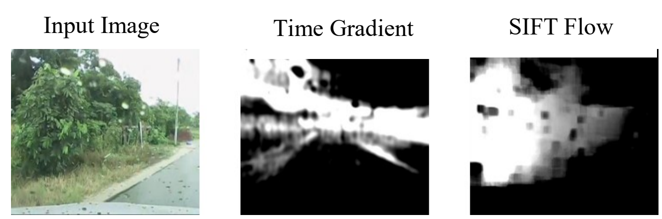

Figure 6. Raindrop features accumulated over 100 frames

We compared our method with Ito et.al.’s method. Our Proposed method and Ito et.al.’s method, detects raindrops in the image/ frame from the single image. One more commonality is usage of MSER to find raindrop candidates. Figure 7 shows results of Ito et.al.’s method and our proposed method for set of 5 different images.

Table 1 summarizes precision and recall rates for the test images shown in figure 7. It’s quite evident from Table 1, the proposed method performs better than Ito et.al.’s implementation. The following procedures contributed to our method’s better performance:

- Adaptive Threshold. Our method adaptively tunes MSER detect parameters according to the image environment. It helps in finding more stable raindrop candidates.
- Edge based detection. This procedure helps in detecting smaller rain drops.
- Filters. Set of additional filters designed on raindrop properties. These additional filters reduce false detection rate.

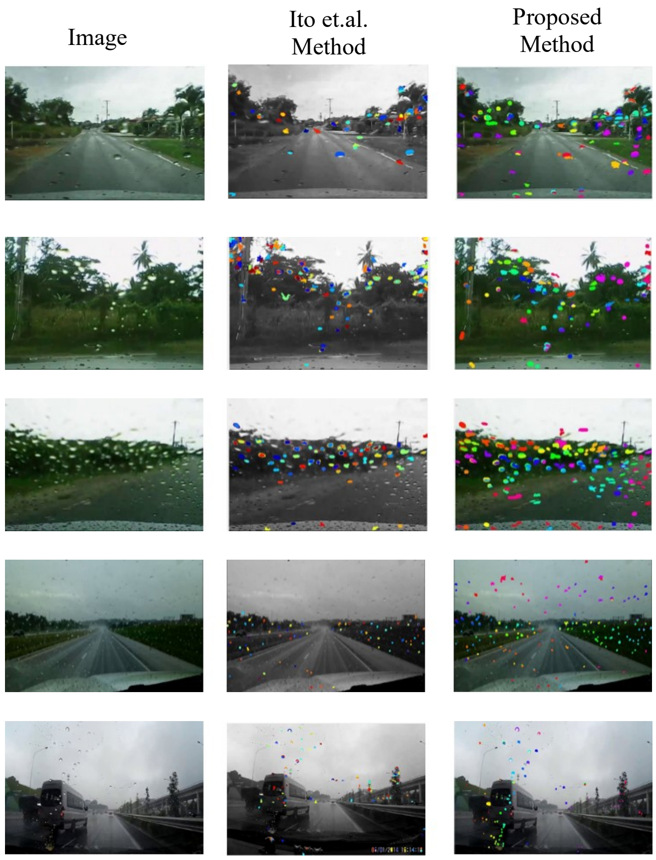

Figure 7. Raindrop detection results using Ito et.al. Approach and proposed approach, where the colored regions indicate the region detected as raindrops (both true and false detections)

	Method	Precision (%)	Recall (%)
Image 1	Ito et.al.'s Method	63.63	51.21
	Proposed Method	70.58	87.80
Image 2	Ito et.al.'s Method	66.66	85.71
	Proposed Method	76.92	95.23
Image 3	Ito et.al.'s Method	96.38	74.07
	Proposed Method	96.07	90.74
Image 4	Ito et.al.'s Method	98.59	38.88
	Proposed Method	98.61	78.88
Image 5	Ito et.al.'s Method	57.14	26.66
	Proposed Method	94.73	60.0

Table 1. Precision and Recall summary for the images shown in figure 7.

Conclusion

Raindrops adherent to a transparent medium (windshield) between a scene and a capture system come in all shape, size and saturation owing to variation in rain density, surface tension of windshield, the distance between windshield and camera, the field of view, environmental lighting etc. However, we can conclude from our results that these drops concentrate rays and form extremal regions in the image. We used MSER to find these regions and applied a series of filters to reduce false positives. Our results showed considerable improvement compared to the nearest work by Ito et.al. Although, we believe a more flexible and adaptive threshold selection method will render the solution more robust. Another modification that can be tested is application of other modified MSER or region detection techniques which are more suitable for tracking in video data [23][24]. A relevant approach to the problem could be from a computational photography point of view [23] or a multi-modal approach which combines information from different forms of capture. One scenario that has not been attempted in literature extensively or by us is raindrop detection at night time, which bring a gamut of associated problems with it.

References

[1] R. Fattal, "Single image dehazing" ACM transactions on graphics (TOG), vol. 27, no. 3, p. 72, 2008.

[2] K. Garg and S. K. Nayar, "Vision and rain" International Journal of Computer Vision, vol. 75, no. 1, p. 3, 2007.

[3] X. Zhang, H. Li, Y. Qi, W. K. Leow, and T. K. Ng, "Rain removal in video by combining temporal and chromatic properties" in Multimedia and Expo, 2006 IEEE International Conference on, pp. 461-464, IEEE, 2006.

[4] S. You, R. T. Tan, R. Kawakami, Y. Mukaigawa, and K. Ikeuchi, "Adherent raindrop modeling, detection and removal in video" IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 9, pp. 1721-1733, 2016.

[5] D. Eigen, D. Krishnan, and R. Fergus, "Restoring an image taken through a window covered with dirt or rain" in Proceedings of the IEEE International Conference on Computer Vision, pp. 633{640, 2013.

[6] R. T. Tan, "Visibility in bad weather from a single image" in Computer Vision and Pattern Recognition, 2008. IEEE Conference on, pp. 1-8, 2008.

[7] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior" IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 12, pp. 2341-2353, 2011.

[8] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks" in European Conference on Computer Vision, pp. 154-169, Springer, 2016.

[9] K. Garg and S. K. Nayar, "When does a camera see rain?" in Tenth IEEE International Conference on Computer Vision, vol. 2, pp. 1067-1074, 2005.

[10] P. C. Barnum, S. Narasimhan, and T. Kanade, "Analysis of rain and snow in frequency space" International journal of computer vision, vol. 86, no. 2, pp. 256-274, 2010.

[11] J.-H. Kim, J.-Y. Sim, and C.-S. Kim, "Video deraining and desnowing using temporal correlation and low-rank matrix completion" IEEE Transactions on Image Processing, vol. 24, no. 9, pp. 2658-2670, 2015.

[12] L.-W. Kang, C.-M. Yu, C.-Y. Lin, and C.-H. Yeh, "Image and video restoration and enhancement via sparse representation" Emerging Technologies in Intelligent Applications for Image and Video Processing, vol. 1, 2016.

[13] V. Santhaseelan and V. K. Asari, "Utilizing local phase information to remove rain from video" International Journal of Computer Vision, vol. 112, no. 1, pp. 71-89, 2015.

[14] J. Gu, R. Ramamoorthi, P. Belhumeur, and S. Nayar, "Removing image artifacts due to dirty camera lenses and thin occluders" in ACM Transactions on Graphics (TOG), vol. 28, p. 144, ACM, 2009.

[15] N. Einecke, H. Gandhi, and J. Deigmoller, "Detection of camera artifacts from camera images" in, IEEE 17th International Conference on Intelligent Transportation Systems (ITSC), pp. 603-610, 2014.

[16] V. Akkala, P. Parikh, B. Mahesh, A. S. Deshmukh, and S. Medasani, "Lens adhering contaminant detection using spatio-temporal blur" in International Conference on Signal Processing and Communications (SPCOM), pp. 1-5, IEEE, 2016.

[17] M. Roser and A. Geiger, "Video-based raindrop detection for improved image registration" in IEEE 12th International Conference on Computer Vision Workshops, pp. 570-577, 2009.

[18] M. Roser, J. Kurz, and A. Geiger, "Realistic modeling of water droplets for monocular adherent raindrop recognition using bezier curves" in Asian Conference on Computer Vision, pp. 235-244, Springer, 2010.

[19] D. D. Webster, "Automatic rain drop detection for improved sensing in automotive computer vision applications" Master's Thesis 2014.

[20] K. Ito, K. Noro, and T. Aoki, "An adherent raindrop detection method using MSER" in Signal and Information Processing Association Annual Summit and Conference (APSIPA), pp. 105-109, IEEE, 2015.

[21] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions" Image and vision computing, vol. 22, no. 10, pp. 761-767, 2004.

[22] T. Xue, M. Rubinstein, C. Liu, and W. T. Freeman, "A computational approach for obstruction-free photography" ACM Transactions on Graphics (TOG), vol. 34, no. 4, p. 79, 2015.

[23] M. Donoser, H. Riemenschneider, and H. Bischof, "Shape guided maximally stable extremal region (mser) tracking" in International Conference on Pattern Recognition (ICPR), pp. 1800-1803, 2010.

[24] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Scha_alitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors" International journal of computer vision, vol. 65, no. 1-2, pp. 43-72, 2005.

Author Biography

Vijay C S *received his BE degree in Telecommunication from M S Ramaiah Institute of Technology, VTU; and MS {Research} in Image Processing and Computer Vision (IPCV) from Indian Institute of Technology, Madras. After graduation he worked with H P Labs India, Samsung R & D Bengaluru. At present he is a PhD Student at National Institute of Advanced Studies, IISc Campus and is working on Electricity system modelling.*

Radhesh Bhat *received his BE degree in Electronics and Communication from Visvesvaraya Technological University in 2004. Currently he is heading digital imaging division at PathPartner Technology Pvt. Ltd. His research interests include camera image signal processing pipeline, image quality, computer vision and deep learning*

Vijaya Ragavan *received his BE degree in Electronics and Communication from University of Madras; and Masters in Electronics engineering from the Anna University, Chennai (2006). Since then he has worked in the Multimedia and ADAS Division at PathPartner Technology Pvt. Ltd. in, Bengaluru. His research interests include algorithms for ADAS on low power heterogeneous systems.*

Free access to this paper is brought to you with the generous support of ON Semiconductor.
All research funding for this paper is referenced in the text; unless noted therein, no research funding was provided by ON.