Classification of Weather Situations on Single Color Images

Martin Roser and Frank Moosmann

Institut für Mess- und Regelungstechnik
Universität Karlsruhe (TH)
D-76131 Karlsruhe, Germany
Email: {roser, moosmann}@mrt.uka.de

Abstract— Present vision based driver assistance systems are designed to perform under good-natured weather conditions. However, limited visibility caused by heavy rain or fog strongly affects vision systems. To improve machine vision in bad weather situations, a reliable detection system is necessary as a ground base. We present an approach that is able to distinguish between multiple weather situations based on the classification of single monocular color images, without any additional assumptions or prior knowledge. The proposed image descriptor clearly outperforms existing descriptors for that task. Experimental results on real traffic images are characterized by high accuracy, efficiency, and versatility with respect to driver assistance systems.

I. INTRODUCTION

Vision based driver assistance systems (DAS) are currently designed to perform under good-natured weather conditions. Unfortunately, limited visibility often occurs in daily life (e.g. heavy rain or fog). As this strongly affects the accuracy or even the general function of vision systems, the actual weather condition is a valuable information for assistance systems. Based on the results of weather classification, specialized approaches for each class can be invoked to improve cognition. This will form a key factor to expand the application of DAS from selected environmental conditions to an overall approach.

Little work has been done on weather related issues for in-vehicle camera systems so far. Nayar [8] studied the visual effects of rain and came up with a photometric rain drop model that describes refraction and reflection of light by a rain drop. Additionally, they determined the effect of camera parameters on image disturbance and developed an approach of detecting and removing rain from videos. Narasimhan [14], [15], [16] analyzed images taken under poor static weather conditions. They used the Koschmieder model to estimate scattering coefficients of the atmosphere and restore the contrast of weather degraded images. Even though Nayar as well as Narasimhan reported impressive results for their specific problems, these approaches can't be easily transferred to automobile applications. All approaches assume a static observer whereas in automobile applications, egomotion of the camera is obviously the normal case. In addition, the numerous shapes of rain drops on a windshield will complicate the employed models significantly. Work on weather related issues in automobile applications

has been conducted by [9] and [10]. Hautiere [9] estimated the visibility distance using charge-coupled device cameras. Kurihata [10] used a machine learning approach with rain-drop templates, so called *eigendrops*, to detect rain drops on windshields. However, both lack a holistic approach to deal with all kinds of adverse weather conditions. In this contribution, we propose a general approach suitable for any kind of weather situation and for any egomotion. As a first step, we present an image classification method that reliably distinguishes between certain weather conditions.

Whereas little work has been done on DAS for bad weather situations, scientific research in image *classification* or *categorization* is very broad. Generally, the goal is to decide whether an image belongs to a certain category or not. Depending on the application, categories can include various natural scenes [11], but often images are tested for the presence of a certain object category, e.g. [5], [13], [17], [19]. All modern approaches are based on the extraction of local image features, as global features turned out to be not robust enough to deal with variations in view, lighting, occlusion and object variations. Different kinds of local features have been proposed with histogram-based features like SIFT [12], HOG [3], and shape context [1] being among the most discriminant. However, these features perform weakly for the intended task. Based on local features, machine learning classification approaches are often proposed that range from simple decision trees [13] up to the introduction of additional semantic layers as the very popular *bag-of-features* approach [2], [19]. While these approaches have achieved remarkable results for generic image categorization, no such system has been proposed yet to distinguish weather situations. Additionally, most existing features are based on grayscale images and only few approaches have tried to use color features, e.g. [20]. We believe that color casts due to atmospheric effects may provide valuable, additional information.

The key contributions of this paper are the development of robust histogram features for the task of weather recognition, and their application in an efficient and effective image classification framework. The method we propose works on single monoscopic color images from in-vehicle cameras and extracts robust and meaningful histogram features, as depicted in section II. In section III, we then apply a support vector machine (SVM) on the
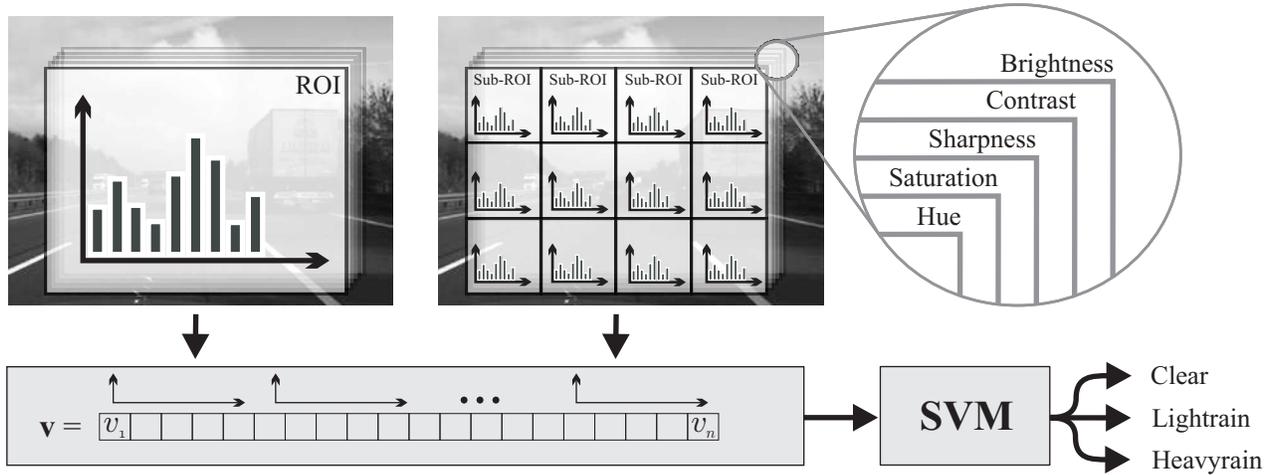
Fig. 1.   Proposed method: Several histogram features (like brightness, contrast, etc.) are calculated in different regions of the image and gathered in one large vector. The vector is then classified by a Support Vector Machine to obtain the estimated weather situation.

features to classify the image into one of the classes $C = \{\text{clear weather}, \text{light rain}, \text{heavy rain}\}$. Section IV shows detailed experimental classification results, discusses the proposed features significance and compares them to standard image descriptors.

II. FEATURE EXTRACTION

A robust weather classification technique is depending on reliable, strong environmental features. Single images, in principle, provide sufficient information for that task.

We first define regions of interest (ROI) which will be used for feature extraction. As depicted in Fig. 1, one global ROI gathers information about the overall effect of weather on the image. In addition, this region is divided into twelve sub-ROI's that cover the local (and distant-dependent) effects in more detail. Within each ROI, several features are evaluated: *local contrast*, *minimum brightness*, *sharpness*, *hue* and *saturation*, detailed below. As some features cannot be computed pixelwise, a ROI is subdivided into 10x10 pixel blocks, and each feature is computed in each block. All features return values between 0 and 1 and vote for one bin of their ROI and feature dependent histogram. Thus, there are 13 ROIs with 5 histograms each.
Histogram bins do not describe the actual local features very accurately, but on the other hand they are very robust in terms of outliers and noise. As we are interested in the overall distribution of features within the image, histogram bins describe the image information appropriately. Beyond that, the quantity of extracted features directly influences the complexity of classification in terms of accuracy, computational time and number of required training images. We discretize the feature histograms into 10 bins as it turned out to be a good compromise between descriptors accuracy and classification effort. The proposed features are presented in greater detail in the following.

A. Contrast and Minimum Brightness

In clear weather conditions, the radiance from a scene point reaches the observer unaltered. However, dealing with adverse weather conditions, atmospheric effects cannot be neglected anymore. In recent literature [14], [15], [16], [9], the *Koschmieder Model* has been established as a description of the atmospheric effects of weather on the observer

$$E = I_\infty \rho e^{-\beta(\lambda)d} + I_\infty(1 - e^{-\beta(\lambda)d}), \qquad (1)$$

where E is the pixel brightness, I_∞ is the background intensity, ρ is the normalized radiance of a scene point [15], d is the distance and $\beta(\lambda)$ is the scattering coefficient. Note that β is a function of the wavelength λ whose relationship is given by Rayleigh's law [14], [16]. For small atmospheric particles like fog or haze, β can be assumed to be constant. Equation (1) implies that the irradiance and thus the brightness observed by each pixel of the sensor is altered by two fundamental scattering phenomena: attenuation and scattered light. In other words, light directly transmitted from the scene point will be exponentially attenuated and superimposed by the environmental illumination that will be refracted towards the observer. For scene points with a low normalized radiance ρ, the direct transmission term can be neglected. Hence, we expect an increasing pixel brightness due to scattered light:

$$E_{\min} \propto (1 - e^{-\beta d}). \qquad (2)$$

In other words, the minimum local pixel brightness will increase with β and d according to the second term in the right-hand side of (1).
The local contrast C can be defined as

$$C = \frac{E_{\max} - E_{\min}}{E_{\max} + E_{\min}}, \qquad (3)$$

where $E_{\max}$ and $E_{\min}$ are the local extrema of the pixel brightness. To increase robustness of contrast estimation, we determine brightness extrema by averaging the darkest and

brightest pixels within the ROI's. Inserting (1) in (3) yields

$$C = \frac{\rho_{\max} - \rho_{\min}}{\rho_{\max} + \rho_{\min} + 2(e^{\beta d} - 1)}. \tag{4}$$

As a result, the local contrast solely depends on scene point properties (which remain constant), distance d and the scattering coefficient β.

B. Sharpness

Clearly distinguishable objects under friendly weather conditions are expected to have sharp edges with large contrast differences. In addition to the contrast feature discussed above, a gradient based method, called the *Tenengrad Criteria* [18], is used to determine the sharpness of the test images. It is based on an average determination of the sobel gradient magnitude

$$T = \frac{\sum_i \sqrt{S_X^2(i) + S_Y^2(i)}}{\sum_i 1}, \tag{5}$$

with $i = [1..\text{all pixels}]$ and the sobel filter responses being S_X, S_Y. This method originates from autofocusing tasks where two images with identical scene information are evaluated according to their sharpness. It fails when applied to images with different context due to the following occurrences: contrast-variance and edge-quantity-variance. Ferzli and Karam [6] proposed a perceptual-based sharpness metric which is invariant to contrast and edge quantity. Similar to their approach, but with slight variations in detecting *edge pixels* and weighting the influence of contrast, we define the *Advanced Tenengrad Criteria* as

$$T_{\mathrm{adv}} = \frac{\sum_i \delta_i \ \rho(i) \sqrt{S_X^2(i) + S_Y^2(i)}}{\sum_i \delta_i}, \tag{6}$$

where $\delta_i = 1$ if pixel i is an *edge pixel* (0 otherwise) and $\rho(i)$ is a weighting factor that is assumed to be inversely proportional to the local contrast.

C. Color Features

Grayscale features are widely used for image processing tasks that range from low level algorithms to highly sophisticated modules, though there is growing attention to color information [20], [7] in feature extraction and tracking topics. Dealing with adverse weather conditions and limited visibility where the significance of features decreases, we attach high importance to additional color information. We extract hue and saturation from the HSV color space. For robustness, local mean values are taken in each 10x10 pixel block.

For all ROIs and features, their values are extracted block-wise and summarized by a histogram. Then, we combine all histograms into one extended descriptor vector, so we get vector $\mathbf{v} = (v_1, ..., v_n)$ with $n = (13 \text{ ROIs}) * (5 \text{ features}) * (10 \text{ bins}) = 650$ scalar elements describing the image.

III. CLASSIFICATION

In this section, we will use the extended descriptor vector $\mathbf{v}$ as described in the last section to decide on the image class. In our case, the classes correspond to weather situations which we divide into $C = \{\text{clear weather}, \text{light rain}, \text{heavy rain}\}$. Thus, the problem of classification can be thought of as finding some function f that maps from descriptor space $\mathcal{D}$ into the classes C with $c = f(\mathbf{v})$, where $f : \mathcal{D} \to C$.

For a descriptor space with a small number of dimensions, such a function f can be designed by hand, whereas for high-dimensional descriptor spaces (e.g. discussed problem: $\mathcal{D} = \mathbb{R}^{650}$) this becomes nearly impossible for a human. The machine learning framework can be used to find such a function from training examples. Numerous methods have been proposed [4] using techniques like k-Nearest-Neighbor, Decision Trees, Neural Networks and Support Vector Machines (SVM).

As SVMs are simple, fast, and powerful, we decided to use them as our learning and classification method. In principle, a linear SVM generates a hyperplane in the descriptor space $\mathcal{D}$ and classifies descriptors by calculating on which side of the hyperplane the descriptor vector (=point) lies. Mathematically, the hyperplane is represented by its normal vector $\mathbf{w}$ with offset b, then for a given descriptor $\mathbf{v}$ a score is calculated by $s = \mathbf{w}^T\mathbf{v} - b$ and the final decision is $(s \geq 0)$. As a hyperplane can only separate two classes, several hyperplanes are needed for the multiclass case, and the scores from each hyperplane have to be combined to get the final classification.

The hyperplane parameters $\mathbf{w}$ and b are optimized in the learning stage to separate the two classes as far as possible. After training, the weights vector $\mathbf{w}$ can be evaluated to get the significance of single features for the classification outcome. Huge values correspond to discriminant features, whereas small values indicate weak features.

One of the advantages of SVM's is that kernel methods can be incorporated in the algorithm. With them, non-linear decision boundaries can be found. We tested two very common kernels, linear and RBF (Radial Basis Functions), with the result that RBF may outperform the linear kernel. However, since one parameter for the RBF kernel has to be optimized manually and our descriptor space is big enough that linear separation is sufficient, we preferred applying a linear kernel.

IV. EXPERIMENTS

In order to overcome the problem of limited image data within the widespread field of adverse weather, we built up a database with video sequences (currently 150 sequences ≈ 500000 single images), labeled according to their particular weather conditions. We randomly selected images from the database to build up our fixed training and testing data sets. We ensured that no image is used for both training and testing at the same time, as well as we ensured equal amount of images for each class.

Fig. 2. Some example images from our image database.

TABLE I

CLASSIFICATION RESULTS FOR SUBSET 1 (EXPRESSWAY ONLY), SUBSET 2 (+ RURAL SCENES) AND SUBSET 3 (+ RURAL AND URBAN SCENES). ROWS CONTAIN THE CORRECT CLASSES (THEY SUM UP TO 180, 300 AND 450 RESPECTIVELY), COLUMNS THEIR CLASSIFICATION RESULTS.

	SUBSET 1			SUBSET 2			SUBSET 3		
	clear	light rain	heavy rain	clear	light rain	heavy rain	clear	light rain	heavy rain
clear	178	2	0	275	24	1	411	39	0
light rain	2	178	0	17	253	30	67	341	42
heavy rain	0	7	173	0	24	276	5	47	398

180 images for each class
Total error rate: 2.04%
(correct: 529, wrong: 11)

300 images for each class
Total error rate: 10.67%
(correct: 804, wrong: 96)

450 images for each class
Total error rate: 14.81%
(correct: 1150, wrong: 200)

A. Classification results for linear SVM

We compose three subsets with increasing demands on the classification by expanding the environmental conditions from expressway only to all possible scenery:

- SUBSET 1: This subset is limited to expressway scenes only with altogether 1080 images.
- SUBSET 2: Here, we expanded the experiments to rural environments, taking 900 expressway scenes and 900 rural scenes into account.
- SUBSET 3: The last subset is similar to SUBSET 2 but with additional 900 images of urban environments.

Fig. 2 shows some example images which illustrate the difficulty of the task. Each category contains images of a large range of brightness and color values, so any single feature would not be sufficient to detect the weather situation with an acceptable confidence.

Experiments reveal that for expressway scenes accurate classifiaction is achieved. We investigate the results in Table I in more detail by applying binary classification to the image sets, that means we only take images from 2 classes. Images that have been classified to the omitted class before, are reassigned to the remaining two classes. It turns out that for subset 1 we achieved an error-free classification between *clear* and *heavy rain* (correct: 360, wrong: 0, error rate: 0%). The error rate between *clear* and *light rain* is 1.1% (correct: 356, wrong: 4, error rate: 1.1%). The most uncertain decision is between *light rain* and *heavy rain* (correct: 353, wrong: 7, error rate: 1.9%). Even humans would not unanimously agree on the correct category of images of these two classes, as the border between light and heavy rain is fluent. Fig. 3(a) shows the corresponding ROC curves, which emphasizes the quality of the classification result.

With increasing demands due to changing environments (subset 2 and subset 3), the accuracy decreases. This is based on the fact that distance dependent features can hardly be extracted from rural scenes since the sub-ROIs do not reflect a robust distance estimation anymore (obstacles in front of the vehicle, closed scene with surrounding objects...). Classwise comparison of the results for subset 2 show error rates up to 10.2%. Remarkably, classification between *clear* weather and *heavy rain* is still very accurate (correct:
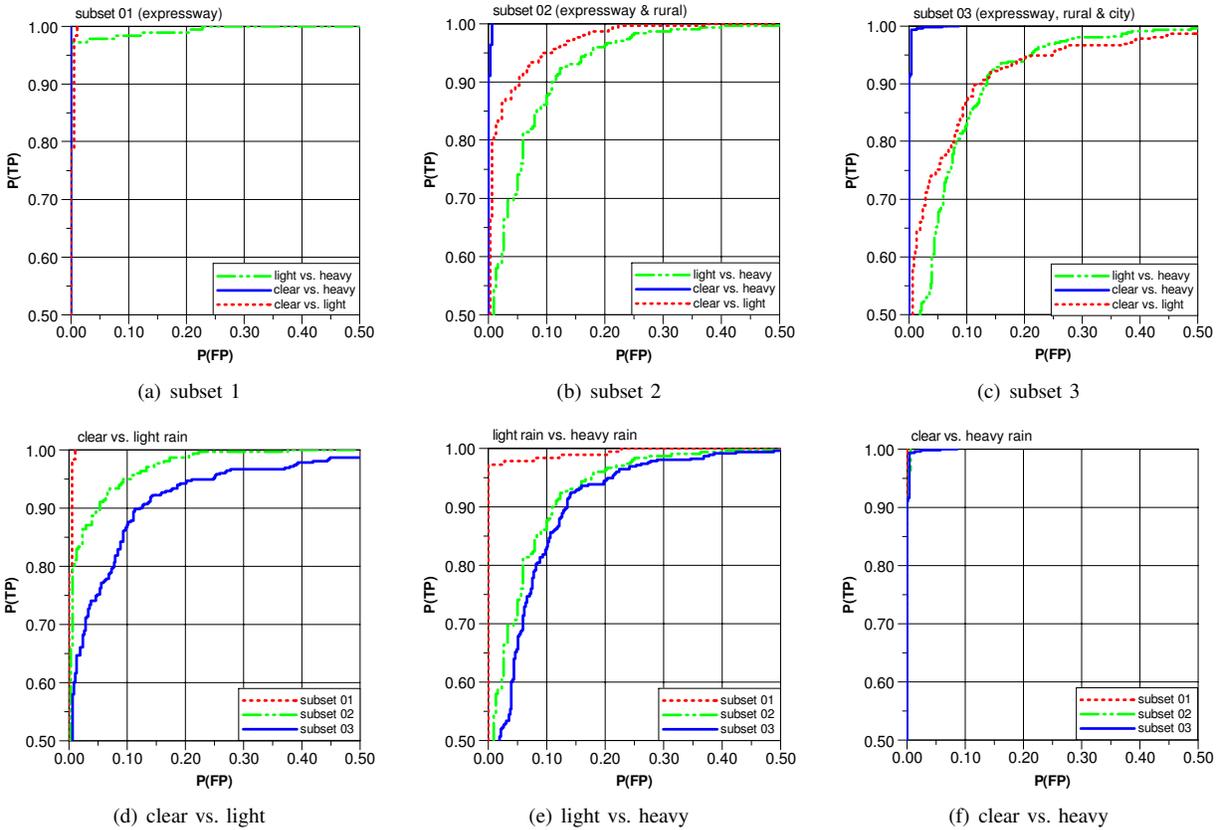
Fig. 3. Receiver Operating Characteristic (ROC) curves for classification results. Axis cropped as to show the top left quarter only.

597, wrong: 3, error rate: 0.5%). In subset 3, classification between *clear* and *heavy rain* again remains at a low error rate of less than 1% (correct: 892, wrong: 8, error rate: 0.9%).

Most misclassifications arise from situations that are hard to define, i.e. that are somewhere inbetween *clear* and *light rain* or *light rain* and *heavy rain*. The second main source of error are outlier images where bridges and other objects confuse the extracted image statistics. However, since weather conditions do not change instantly, it is possible to classify multiple times and combine the results to improve accuracy. Further optimization is possible by using non-linear SVM that better suit the application.

Execution times are $\approx 1.8s$ per image on a Centrino 2.4GHz running Matlab. Nearly all time is spent on feature extraction, which can be reduced significantly by using optimized code (possibly on the GPU for real time usage). Anyway, since high measuring rates > 0.5Hz are not necessary, this approach is already applicable for DAS.

B. Feature evaluation

To benchmark our approach compared to existing methods, the proposed features are evaluated in regard to their significance for the classification decision as well as their overall performance. However, studying the effects of omitted features on classification results leads to a known problem with linear SVM-kernels: If the dimensionality of

descriptor space $\mathcal{D}$ drops below a lower bound, a linear feature separation cannot be drawn anymore. For that reason, we used a non-linear RBF kernel for subsequent evaluation.

In section II, we proposed a novel image descriptor for the task of weather classification. In oder to benchmark its performance, we additionally extracted color wavelets as well as a combination of SIFT features and color histograms and compared the classification results. As depicted in Fig. 4(a), the proposed features clearly outperforms both standard image descriptors.

Low error rates in SVM classification can only be achieved with optimal feature selection. As mentioned in section III, parameter $\mathbf{w}$ of the SVM tells us the significance of each dimension of descriptor space $\mathcal{D}$. In our experiments, all feature weights are evenly distributed, that means not one feature alone or any combination of some features is able to achive high discrimination, the descriptiveness lies in the combination of all proposed features. We verified these results by omitting single features and running tests again. Altogether, there are $\sum_{i=1}^{5} \binom{5}{i} = 31$ possible combinations of the proposed features. Fig. 4(b) shows the classification error for all possible feature combinations. It can be observed, that all classification errors considering the same amount are close to their mean, whereas a classification improvement can only be achieved by taking additional features into account.
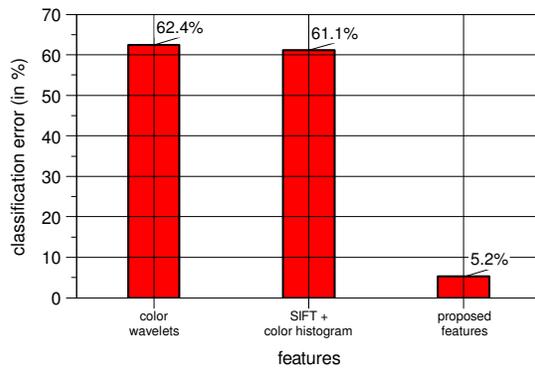
(a) Feature benchmark (b) Feature combination
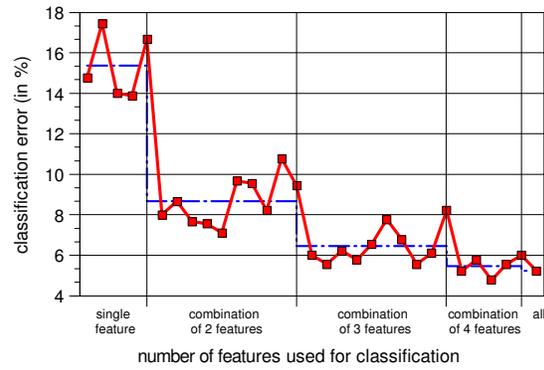
Fig. 4. Feature evaluation

V. CONCLUSIONS AND FUTURE WORK

We presented a novel approach to weather conditions classification based on single monoscopic color images. Thereby we did not make any assumptions like having a static camera nor is the approach restricted to a specific application setting. The approach will work without restrictions in terms of camera, frame rate or environment. The approach achieves low error rates of less than 1% for the distinction between clear weather and heavy rain and even acceptable error rates for the three-class-case. We discussed that the main misclassifications can be reduced significantly by classifying consecutively every few seconds and filtering outliers, as the weather situation cannot change very rapidly. It could be shown that all proposed features are important to the same degree for a robust weather classification and clearly outperform standard image descriptors.

Future work will expand C by adding other weather situations like fog to our database. Improvements of the overall classification results could be achieved by further in-depth studies to non linear SVM-kernels. Specialized methods on certain weather situations can then be invoked based on the classification result to improve existing vision algorithms.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *NIPS*, pages 831–837, 2000.

[2] G. Csurka, C. Dance, L. Fan, J. Williamowski, and C. Bray. Visual categorization with bags of keypoints. In *ECCV'04 workshop on Statistical Learning in CV*, pages 59–74, 2004.

[3] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *International Conference on Computer Vision & Pattern Recognition*, volume 2, pages 886–893, INRIA Rhône-Alpes, ZIRST-655, av. de l'Europe, Montbonnot-38334, June 2005.

[4] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern classification*. Wiley, 2. ed. edition, 2001.

[5] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman. Learning object categories from google"s image search. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision*, pages 1816–1823, Washington, DC, USA, 2005. IEEE Computer Society.

[6] R. Ferzli and L.J. Karam. Human visual system based no-reference objective image sharpness metric. In *IEEE Internatl. Conference on Image Processing*, pages 2949–2952, Atlanta, GA, 2006.

[7] P. Gabriel, J.-B. Hayet, J. Piatera, and J. Verly. Object tracking using color interest points. In *IEEE Conference on Advanced Video and Signal Based Surveillance*, pages 159–164, 2005.

[8] K. Garg and S.K. Nayar. Vision and rain. *Internatl. Journal of Computer Vision*, 75(1):3–27, 2007.

[9] N. Hautière, R. Labayrade, and D. Aubert. Real-time disparity contrast combination for onboard estimation of the visibility distance. In *IEEE Transactions on Intelligent Transportation Systems*, volume 7, 2006.

[10] H. Kurihata, T. Takahashi, I. Ide, Y. Mekade, H. Muraseand Y. Tamatsu, and T. Miyahara. Rainy weather recognition from in-vehicle camera images for driver assistance. In *IEEE Proceedings, Intelligent Vehicles Symposium*, pages 205–210, 2005.

[11] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2:2169–2178, 2006.

[12] David G. Lowe. Object recognition from local scale-invariant features. In *Proc. of the International Conference on Computer Vision ICCV, Corfu*, pages 1150–1157, 1999.

[13] Raphaël Marée, Pierre Geurts, Justus Piater, and Louis Wehenkel. Random subwindows for robust image classification. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, volume 1, pages 34–40. IEEE, June 2005.

[14] S.G. Narasimhan and S.K. Nayar. Chromatic framework for vision in bad weather. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 598–605, 2000.

[15] S.G. Narasimhan and S.K. Nayar. Removing weather effects from monochrome images. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 186–193, 2001.

[16] S.G. Narasimhan and S.K. Nayar. Vision and the atmosphere. *Internatl. Journal of Computer Vision*, 48(3):233–254, 2002.

[17] Andreas Opelt and Axel Pinz. Object localization with boosting and weak supervision for generic object recognition. In *SCIA*, pages 862–871, 2005.

[18] J.L. Pech-Pacheco, G. Cristóbal, J. Chamorro-Martínez, and J. Fernández-Valdivia. Diatom autofocusing in brightfield microscopy: a comparitive study. In *IEEE 15th Internatl. Conference on Pattern Recognition*, volume 3, pages 314–317, 2000.

[19] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *ICCV*, volume 2, pages 1470–1477, October 2003.

[20] J. Weijer and C. Schmid. Coloring local feature extraction. In *Lecture Notes in Computer Science*, Berlin, Germany, 2006. Springer.