https://doi.org/10.2352/ISSN.2470-1173.2017.12.IQSP-229

A RGB/NIR Data Set For Evaluating Dehazing Algorithms

Julia Lüthen, Arnold & Richter Cinetechnik, Munich, Germany
Julian Wörmann, Martin Kleinsteuber, Department of Electrical and Computer Engineering, TU München, Munich, Germany
Johannes Steurer, Arnold & Richter Cinetechnik, Munich, Germany

Abstract

Realistic test data is needed to evaluate and rank the performance and quality of dehazing algorithms for image enhancement. Especially for professional photography and cinematography, this test data has to fulfill high quality standards including high dynamic range, sufficient resolution and natural color reproduction. For this purpose, we present a new multispectral data set that includes RGB and near-infrared (NIR) images captured by two professional digital motion picture cameras. Compared to existing data sets, the benefits of our set are threefold. Due to our two camera setting we are able to provide synchronous and well registered RGB/NIR image pairs captured at the same instant of time. High quality real image sequences allow future algorithms to take account of the temporal consistency of the dehazed output images. Furthermore, to facilitate a uniform and fair evaluation of different algorithms we provide ground truth images for selected RGB/NIR image pairs. The data set is freely available at http://www.arri.com/innovations/.

Introduction

Haze in outdoor images is a weather phenomenon that leads to loss of contrast, visibility and saturation and that is hardly to avoid while capturing the image. Even in almost clear weather situations there will always be decreasing visibility and contrast in long distances (Fig.1).

In the last decade, image dehazing has become a popular research area resulting in various algorithms that aim at removing the haze out of haze-afflicted images. Nowadays, these algorithms have even found their way in image processing software like Adobe Photoshop, allowing the user to easily remove fog or haze from a photograph. Ongoing from algorithms that rely on a single input image [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] to estimate the missing parameters to recover contrast, visibility and saturation, approaches that utilize additional input images that provide further information about the captured scene have emerged recently [13, 14, 15, 16, 17]. In particular near-infrared (NIR) images as additional input have been proven useful to enhance visibility and contrast, due to the fact that haze is a visible phenomenon that decreases with increasing wavelength of the captured spectrum [18].

Regarding this multiple input image setting, a common cause for disturbing artifacts in the dehazed images is the deregistration between the two modalities due to sequentially captured images. Especially images and sequences that include moving objects are prone to these errors. Another drawback of most of the existing real-world data sets is the absence of ground truth. Thus, a performance comparison between different approaches is hard since the only way for an evaluation is to perform user perception studies

Figure 1: Haze in images leads to loss of contrast, saturation and visibility

which are costly and time consuming. Since it is mandatory for user perception tests to evaluate different algorithms at the same instant in time, the evaluation with satisfying ground truth images is more flexible. One approach to circumvent these drawbacks is to synthetically generate the image data as well as the hazy input. However, solely relying on computer-generated haze images often results in parameter settings that are not applicable to real world scenarios.

As a consequence, a data set for an adequate evaluation of dehazing algorithms should offer the possibility to render high quality real-world images and to compare the results with a corresponding ground truth. Our new RGB/NIR data set addresses all these points. Compared to existing data sets the advantages of our set are:

1. Due to our camera setup, we are able to provide well registered high quality RGB/NIR image pairs captured at the same instant of time.
2. Multispectral sequences are made available, allowing the evaluation of multispectral algorithms in terms of temporal consistency, which is not feasible by any other data set.
3. For the first time, it is possible to compare the dehazed images with reference haze free multispectral image pairs, leading to a fair evaluation of single image and multispectral image input algorithms.

The rest of the paper is organized as follows. After the introduction of the physical haze model, which the most significant dehazing algorithms are based on, we summarize existing data sets and dehazing algorithms. Afterwards, the data set generation and the RGB/NIR data set is described in detail. An evaluation of various dehazing algorithms is presented in the evaluation section.

Physical Haze Model

To generate a data set for dehazing algorithms, it is important to understand the characteristics of light scattering, its influence on radiant light and therefore on the camera capture. Most of the existing dehazing algorithms are premised on a haze model to determine the interfering parameters, which we will introduce in the following. A single light ray never finds its way through the atmosphere without being scattered by particles in the air. In general, this phenomenon is described by Mie scattering of electromagnetic waves. Dependent on the particle types and sizes, different weather situations like e.g. haze, fog or rain can be categorized. Haze is an aerosol with particle sizes of $10^{-2} - 1\mu m$ [6]. Since the spectrum of visible and near-infrared light is composed of wavelengths in the range of $\lambda = 380 - 1200nm$, this setting allows to apply the Rayleigh approximation of Mie scattering. According to [18], for particle sizes smaller than one-tenth of the wavelength, the scattering follows Rayleigh's law where the scattering is inversely proportional to the fourth power of the wavelength, i.e.

$$\beta \propto \frac{1}{\lambda^4}, \tag{1}$$

with β as the scattering coefficient and and λ denoting the wavelength. Due to the long wavelengths of near-infrared light, from (1) it is immediately clear that NIR images are less affected by haze and thus are beneficial in recovering structures.

Nayar and Narasimhan [6, 5, 19] split the mechanisms of scattering into two fundamental phenomena. First, a light beam traveling through the atmosphere gets attenuated due to multiple scattering in the medium. Formally, with L_0 denoting the intrinsic luminance and d the scene depth, the *attenuation* is modeled via $L_0 e^{-\beta d}$. Second, the effect that more distant objects from an observer appear lighter than nearer objects, caused by scattering of environmental light, is described in terms of *airlight*. This term is given by $L_s(1 - e^{-\beta d})$, with L_s representing the global atmospheric light. Finally, these two components lead to the widely used haze image formation model [3] that describes the observed intensity L at the pixel position (u, v) via

$$L(u,v) = L_0(u,v)e^{-\beta d(u,v)} + L_s(1 - e^{-\beta d(u,v)}). \tag{2}$$

This formation model provides the basis for state-of-the-art dehazing algorithms presented in the next section.

Related Work

Research on image dehazing concentrates on two different kinds of algorithms: single image input and multiple image input. The former approaches only rely on a single degraded input image to estimate missing information of the captured scene and to reconstruct e.g. contrast and color information. The second group of dehazing algorithms uses image fusion of (multispectral) image pairs to enhance the contrast and details of the resulting image.

Image Dehazing with Single Image Input

Based on a hazy image as single input image, Tan [1] obtains the (global) atmospheric light by analyzing the pixel intensity of the input image and models the airlight via contrast maximization and a cost function using Markov random fields. Fattal [2] estimates the medium transmission and the albedo of the scene to get two locally uncorrelated shading and transmission functions to recover a dehazed image. Single image dehazing formulated as a particular filtering problem is proposed by Tarel et al. [11]. The algorithm focusses on execution speed and thus only relies on simple median filter operations. Kratz and Nishino [12] tackle the image dehazing problem from a probabilistic point of view. Based on a Factorial Markov Random Field of the image, the depth and the albedo of the scene are estimated simultaneously. This formulation allows to impose structural constraints on both the depth and albedo values. This strategy is further pursued in [10]. He et al. [3] propose a dark channel prior to remove haze from a single image input. This model is based on the theory that at least in one color channel of natural haze free outdoor images (and except sky regions), the intensity value of every pixel is near zero. Therefore the intensity of the dark channel of hazy images could be assumed as thickness of haze in these images. This knowledge leads to an estimation of atmospheric light and transmission. Based on the assumptions of the dark channel prior of [3], Meng et al. [4] derive a boundary constraint map by applying a morphological closing on a patch wise transmission. In combination with a weighted L_1-norm based contextual regularization of the single patches the scene transmission is estimated. Additional to the dark channel, Tang et al. [7] extract three other haze relevant features: local contrast, hue disparity and local saturation and assume that these features show strong correlations with the amount of haze. These different correlations help to estimate the transmission more precisely than [3] if the haze-free image content is very bright by nature where the single dark channel prior would fail. Ancuti and Ancuti [8] derive two input images from a single image input. The first input image is a white balanced version of the hazy input image and a second input image is generated by enhancing the contrast of the white balanced input. Additionally, three different weighting maps (luminance, coloration and saliency) are calculated from the two derived input images. The final image is generated by the Laplacian pyramid representation of the input images and the Gaussian pyramid representation of the combined weighting maps. In [9], Fattal recovers the scene transmission by analyzing the color lines of small pixel patches of a hazy input image and validating the result of each pixel patch with a defined formation model. A further interpolation and regularization of the local transmission for the entire input image leads to a dehazed version of the input image.

Image Dehazing with Multiple Image Input

As mentioned before there is also the opportunity to get missing details out of more than one captured image. Schechner et al. [14] use at least two images taken with different polarization orientations to recover the unknown scene radiance and airlight for building a haze model. With a NIR image as additional input, Schaul et al. [15] convert the 3-channel RGB image into one luminance and two chrominance channels and merge the luminance image with the NIR image via edge-aware multiscale representation. This method adds missing details from the NIR image to the RGB but does not recover any missing color in a hazy image. Another RGB/NIR image fusion is done by Zhang and Wang [16], using the dark channel prior to estimate the scene depth and to apply different color transfer schemes to merge the RGB and NIR images. Feng et al. [13] capture a multispectral image pair to estimate the airlight out of the dissimilarities between the two different captured spectra of visible light and near-infrared light. The method of Connah et al. [17] expands the RGB

and NIR image to a higher dimensional structure tensor representation. This representation is mapped to the gradient field of the input RGB image to seek for a reintegrated output image, whose gradient field is as close as possible to the input image.

Data Sets

In 2002, Narasimhan et al. [20] published the *Weather and Illumination Database* (WILD). This data set contains images of one outdoor scene captured automatically each hour over several month to provide different weather and illumination situations. The data set includes ground truth information such as detailed weather information and a corresponding depth map. However, the set is restricted to only one static scene without moving objects. The spatial resolution of 1.5 MPix is only moderate and the frame rate is very low. Brown and Süsstrunk [21] provide 477 different natural RGB/NIR image pairs in different categories (Fig.2). For each scene, the RGB and NIR images were captured sequentially with a conventional DSLR camera where the hot mirror has been removed. First, an infrared blocking filter has been mounted in front of the lens to capture the RGB image. Afterwards, a visible light blocking filter has been used for the NIR image. Due to the sequential capturing, the RGB and NIR images show slight differences (e.g. moving clouds, water, people crossing, etc). An image fusion of these image pairs is prone to artifacts due to the moving image content. Obviously, this acquisition strategy is not applicable to capture image sequences. The Foggy Road Image Database (FRIDA) of Tarel et al. [22] consists of 90 synthetic images in urban road scenes. Each scene set contains one haze free image (ground truth), four images with different haze grades and a depth map (Fig.3). Since the images are computer generated they do not represent natural surfaces and structures like leaves, clouds, wood and stony surfaces and do not show typical sensor characteristics such as sensor noise.

Contributions and Limitations

The goal of our work is to provide a data set that addresses the limitations of previous sets. Thereby, our data set focuses on the evaluation of algorithms that (i) rely on single and (ii) multispectral image input. For this purpose, we provide high quality test images containing different sceneries with varying surfaces, distances, lighting conditions and with natural and artificially generated haze. The artificial haze allows us to additionally provide ground truth images of the respective scene. Regarding the evaluation of multiple input algorithms, we further provide synchronized and well-registered RGB/NIR image pairs captured with our innovative two-camera setup. The synchronized cameras allowed to record interior as well as exterior image sequences and camera panning without obtaining temporal artifacts between the RGB and NIR images caused by moving objects.

Limitations The ground truth data for an objective evaluation and comparison between single image and multiple image input algorithms is only available for the indoor scenes that contain artificially generated haze. The pursued strategy of capturing a hazy image with its corresponding ground truth is restricted to stills and does not allow to provide ground truth for image sequences. Furthermore, the data set does not contain any depth map or images with different polarization orientations, which might serve as additional information in the dehazing process.

Specifications Our data set consists of 11 pairs of stills including 7 outdoor scenes and 4 indoor scenes. Ground truth is available for indoor scenes as well. Another 6 temporal sequences each containing 200 frames (8sec) are provided. Every image has a resolution of 2800x1575 and comprises a bit depth of 16. The RGB images are converted to ITU Rec 709 video-colorspace [23] and the NIR images are converted with a corresponding video-Look-Up-Table. The depth of the indoor scenes is 12 meters on average.

Data Set Generation

A missing feature of existing image data sets like [21] is the simultaneous capturing of the RGB and NIR data in stills and sequences. This feature is required for further image fusion to avoid image artifacts due to any change in the pictures because of moving objects like clouds in the sky, etc. In the following the used equipment and the simultaneous capturing method is explained in more detail.

RGB- and NIR-Camera

For the image capture we used two different configurations of the ARRI Alexa, a professional digital motion picture camera. Since the presented data set is comprised of images captured with the same sensor as many international film productions, it emphasizes the importance of the evaluation of different algorithms based on this data set. Algorithms which produce artifacts on these images would be useless for cinematic post production.

Alexa RGB Sensor The sensor which is built in a standard ARRI Alexa is a CMOS sensor in the APS-C format class with a Bayer pattern color filter array. The camera captures the RGB-signal of the data set using the 16:9-ARRIRAW-format. Since a plain CMOS sensor would be sensitive to wavelengths up to 1200nm there is a lowpass filter pack in front of the sensor to cut the sensitivity to visible light. The filterpack contains an optical lowpass filter as well as an UV- and an IR-filter. This so-called hot mirror is the standard IR-elimination filter of the ARRI Alexa camera.

Alexa NIR Sensor The NIR-images are captured with a special ARRI Alexa camera called Alexa B&W (Black and White). This camera type contains the same CMOS sensor as the standard ARRI Alexa, but without Bayer pattern color filter and an infrared pass (visible light block filter) instead of the hot mirror in front of the lowpass filterpack. All other technical specifications are corresponding to any standard ARRI Alexa.

ARRIRAW format Both images, RGB and NIR, are captured in ARRIRAW-format, since it is the only supported format of the Alexa B&W. Like any other raw-format, an ARRIRAW-file is a one-channel image which is not debayered before the recording. In postprocessing the RGB-images need to be rendered to a 3-channel-image. Since the ARRI Alexa B&W does not contain a Bayer pattern on the sensor, there is no need for any debayering in the postprocessing. Thus, the NIR-image is stored as a one-channel luminance image.

Optics To get the best matching image pair in RGB and NIR, the axial aberration of the NIR-image has to be compensated, since the NIR wavelength leads to a different refraction index and

Figure 2: Example image pair of RGB and NIR capture in [21]

Figure 3: Example images of FRIDA with 4 different fog grades [22]

hence to a slightly different focal length. The image plane of the NIR wavelength was corrected by modifying the flange focal distance (FFD) between the NIR camera sensor and the lens mount. To calibrate the FFD, a standard prime lens was mounted in front of a movable camera sensor that is sensitive to both, visible and IR light. For different axial positions of the sensor the modulation transfer function (MTF) was measured with two different passfilters (visible light and NIR) mounted in front of the lens. The offset between both maxima determines the compensation of the FFD for both modalities.

Synchronous Capture To enable a synchronous capture of the RGB and NIR images, the cameras were mounted on a 3D-mirror rig and the camera sensors were set on sensor sync, meaning that the image frames of both cameras were captured at identical time instances. We used a stereoscopic analyzer [24] for a visual as well as a feature based inspection to correct possible parallax errors due to vertical or horizontal displacements. This setup allows capturing RGB and NIR-image sequences at the same time and from the same viewpoint in the RGB and NIR-camera channels.

Artificial Haze

We used a haze machine and a water-based haze fluid to capture hazy images and their corresponding target image (ground truth). In contrast to, e.g. [7], who synthesize haze on *already existing* images which are randomly collected from the internet we get realistic multiple scattering from natural light sources (atmospheric light) by adding the haze to the scene *before* the image capture. We have reasons to believe that for the artificial haze scenarios, Rayleigh scattering still holds true because the particle size is small enough. In accordance with equation (1) the near-infrared wavelengths are less scattered and thus the NIR images show more details than their corresponding RGB images, which can be easily seen in figure 4. Since the haze fluid affects the scattering of light sources like in real haze images this procedure allows to create almost natural haze scenes.

Ground Truth

In comparison to [21], we offer RGB/NIR image pairs with and without haze (Fig.4). The clean image pairs serve as ground truth and enable an objective comparison between processed images and the clean target images. To generate the ground truth image pairs, we first shot the clean scene to get the target image. After that, the entire scenery was filled with haze to achieve a natural scattered light effect like in the atmosphere. This process took approximately 5-10 minutes. We had another 10 minutes to capture the hazy scene until all the haze was faded away. Ground truth images are available for 4 different indoor scenes.

RGB/NIR Dataset

In this section the different parts of the data set are described in more detail along with potential challenging features for existing and future algorithms. Figures 4,5, and 6 illustrate the diversity of the scenes.

Stills

Ballroom (Fig.4:1,2,5,6) The *Ballroom*-images were captured in front of a white wall. While image fusion techniques might be suitable to enhance the contrast in the RGB images, this image set should be pretty challenging to algorithms that work with a dark channel prior, since there is no dark background in these images. All images have been captured with natural daylight only, appearing through windows from camera left.

Innolab (Fig.4:3,4) The recovery of the original colors, hue and saturation is a very challenging task in image dehazing. Dehazing

algorithms that ignore color management, e.g. by solely merging the converted luminance channel of the RGB input with the NIR input, could generate tinged results. We think that this image pair facilitates the evaluation of the color reproduction ability of the respective dehazing method.

Greenhouse (Fig.4:7,8) The *Greenhouse*-set was captured in a greenhouse to benefit from direct sunlight. The internal setting allows to provide hazy image pairs with corresponding ground truth images. The scene consists of many repetitions like the lighting set or the trusses of the greenhouse as well as many natural surfaces like different kind of plants, wood and the windows of the greenhouse. This image set could help to enhance dehazing algorithms based on RGB- and NIR-image fusion. Algorithms based on combining the details of the RGB- and NIR-image only may have difficulties with the leaves of the plants in the setting because the NIR-image reflects a brighter surface in these parts than the RGB-image.

Natural (Fig.5) To provide some natural sceneries as well, we also included a data set which contains outdoor daylight images without corresponding ground truth. This set includes mountain panoramas, urban architecture and different structures and surfaces like rooftops, concrete, dressed stone, moss and green.

Sequences (Fig.6)

Additionally, we provide six sequences that enable to evaluate the consistency of dehazing algorithms across successive image frames. The sequences are denoted as *Bicycle Ride* (daylight, indoor, artificial haze), *Rhythmic Gymnastics* (daylight, indoor, artificial haze, with 3 different backgrounds), *Mountain Pan*, and *Mountain Cable Car*. For video processing it is not only important to get high-quality processed single frames but it is even more important that rendered image sequences do not show temporal artifacts. As we captured both channels simultaneously we can provide image sequences with camera panning or movements in front of the camera to evaluate algorithms spatially and temporally as well. In the *Rhythmic Gymnastics*-scenes we provide different background colors (white, green and black). Since these sequences include movements, there is no ground truth available.

Evaluation

In this section, we exemplarily present results of various dehazing algorithms to show the advantage of a reference based objective evaluation. For this purpose, we present the results produced by the algorithms of Nishino et al. [10], Meng et al. [4], and Tarel & Hautière [11] which are all publicly available. The parameters are set to the ones suggested by the authors. Furthermore, for selected images we have included results from Tang et al. [7] and the algorithm of Connah et al. [17].

Regarding previous image dehazing publications, the restoration quality of different algorithms is often assessed only visually which is a highly subjective evaluation criteria. In [8], the authors mention that the only existing method for a quantitative interpretation of the dehazing performance is a blind measure introduced in [25] and also used in [11]. In [25], the authors present blind quality metrics that base on the gradient of visible edges in the image before and after contrast restoration. First, the value e denotes the ratio of newly visible edges, i.e. positive values reveal the ability of the dehazing method to restore edges which were not visible in the input image but are existent in the processed one. Since only the plain number of edge pixels is evaluated, this measure can not distinguish between true edges that become visible after dehazing and artificial edges that are generated by the restoration process, e.g. noise. Second, the ratio of the image gradient norms between the restored and the hazy image is computed. An increase in the gradient norm indicates a contrast stretch. The indicator $\bar{r}$ specifies this average gain of visibility level. Important to note is that both measures only operate on gray-value images and thus do not take into account any color information. Since most of the considered algorithms are tested on images that have a resolution comparable to the VGA standard, all images have been resized to 800×450 pixels and converted to a bit depth of 8-bit before processing. Figure 7 depicts the results and blind quality measures of different dehazing methods applied to one of the images from the *Natural* data set. Please note that for these images no ground truth is available. One can observe that all methods are able to increase the contrast and to recover structures that are not visible in the hazy input images. The given quality scores are in accordance with these observations.

Since the presented dataset additionally consists of selected image pairs with ground truth, we are able to apply several full reference based quality metrics. To be precise, we list the widely used mean structural similarity (SSIM) [26] and visual information fidelity (VIF) [27] measures. To take into account the color information we also computed the feature similarity index including the chrominance information (FSIMc) proposed by [28]. Additionally, we applied the color-image-difference (CID) metric, which was introduced in [29]. The outcome of all these measures lies in the interval between zero and one. However, please note that while SSIM, VIF, and FSIMc values close to one indicate perfect reconstruction, a CID value near zero represents a perfect match. In figure 8 the restored images as well as the corresponding values of the quality metrics are presented. The SSIM (except for one of the Innolab images) as well as the VIF values are higher compared to the values computed on the hazy input. This indicates that the restored images contain more structures resulting in an increase in visual information. The positive values of e and the $\bar{r}$ values greater than one further indicate that the output images contain more edge pixels and that the contrast of the edges is enhanced, respectively. However, the range of both values is not consistent across the two different input images. Especially the e values corresponding to the restored *Innolab*-images differ in a great extend from the values achieved after processing the *Greenhouse*-image. Thus, it is very hard to quantify which value range corresponds to an acceptable outcome. The authors of [8] already mention that increasing the local contrast too strong will result in higher values of the indicator $\bar{r}$. As a result, a moderate value for $\bar{r}$ seems to be best which is hard to determine in practice. On the contrary, the reference based quality measures allow for a consistent evaluation across several input images. The mean ranking across the four different metrics reflects the visual perception of the restored images.

Another aspect one can deduce from figure 7 and 8 is that although all methods recover missing structures, the color reproduction seems to be one of the major difficulties in image dehazing. Quantitatively, this can be seen for example in the last row of figure 8, where high values of SSIM and VIF indicate the best

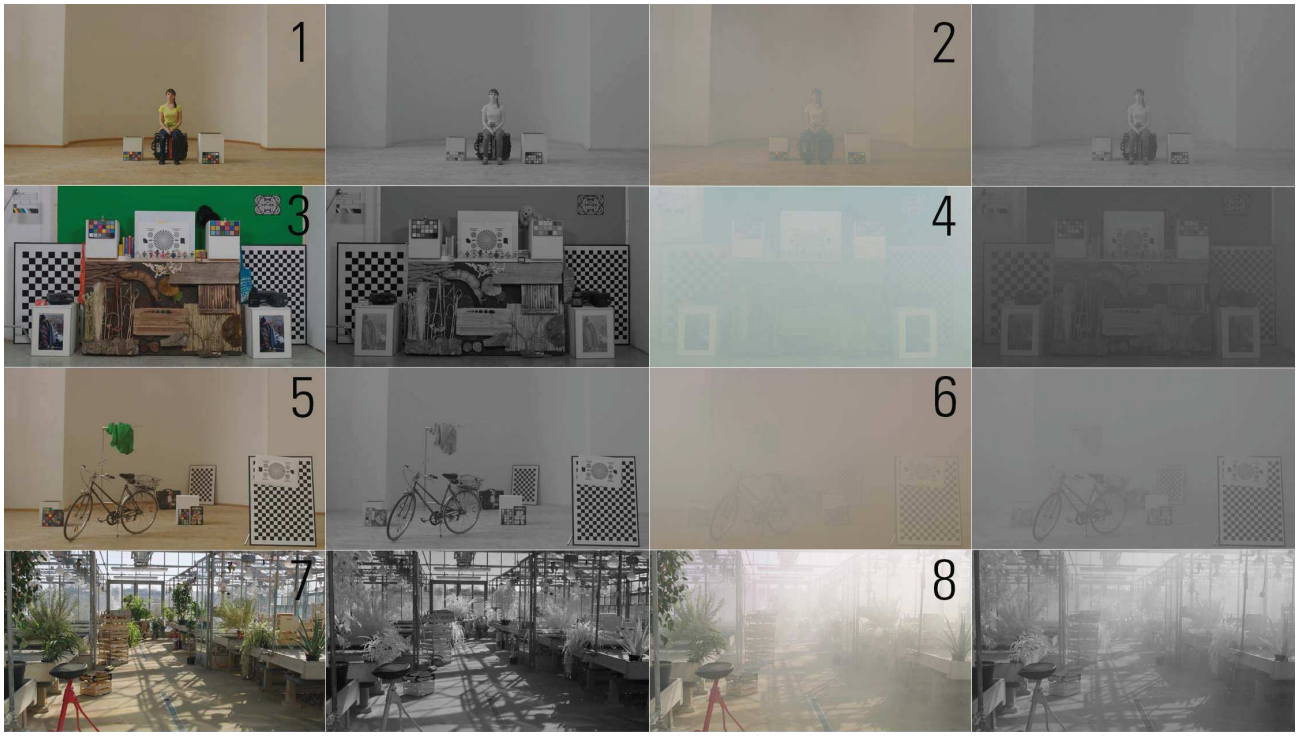


Figure 4: The RGB/NIR data set provides hazy RGB and NIR image pairs (2,4,6,8) with corresponding ground truth image pairs (1,3,5,7) for visual and mathematical comparison

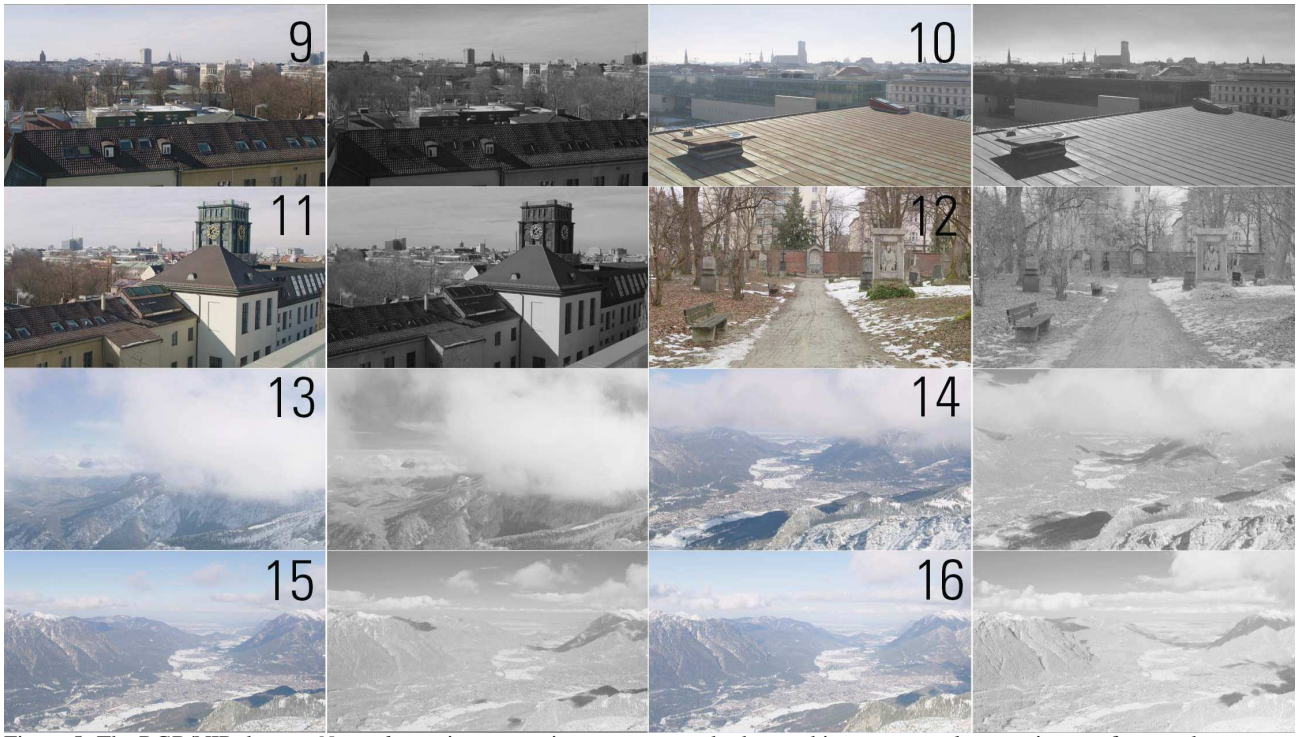
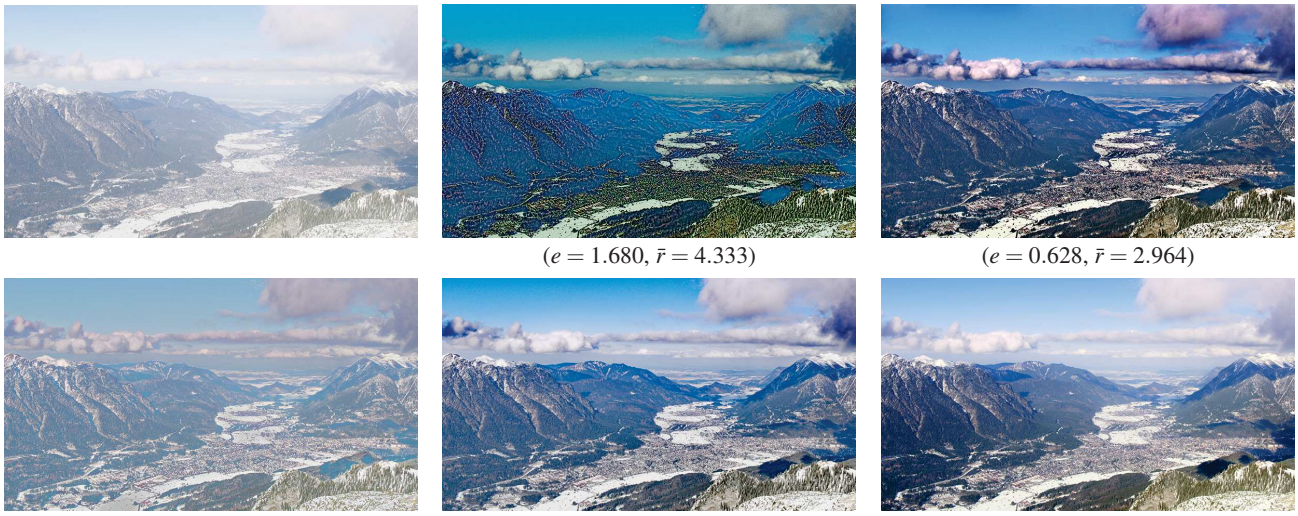


Figure 5: The RGB/NIR data set *Natural* contains mountain panoramas and urban architecture to evaluate various surfaces and structures

reconstruction of image structures, but the tinged outcome results in a large color difference CID.

To conclude, we believe that reference ground truth images will definitely simplify and encourage the development of new approaches that can face the problem of accurate color reproduction. Reference based quality metrics can be calculated quickly, they are easy to interpret and their outcome is in accordance with the visual perception.

Figure 6: Synchronous capture of RGB and NIR sequences for the evaluation of algorithms with multiple image input


($e = 1.680$, $\bar{r} = 4.333$) ($e = 0.628$, $\bar{r} = 2.964$)

($e = 0.780$, $\bar{r} = 1.876$) ($e = 0.437$, $\bar{r} = 2.523$) ($e = 0.359$, $\bar{r} = 1.877$)

Figure 7: Restoration results of different dehazing methods for an image from the *Natural* dataset. Top row, from left to right: Hazy input image. Result of Nishino et al. [10]. Result of Meng et al. [4]. Second row from left to right: Result of Tarel & Hautière [11]. Result of Tang et al. [7]. Result of Connah et al. [17].

Conclusion

In this paper we present a new multispectral RGB/NIR data set which meets the image quality level of professional motion picture imaging in respect of resolution, dynamic range and color reproduction. The data set offers the ability to evaluate and compare single as well as multiple input dehazing algorithms, due to the synchronously captured NIR images. Furthermore, the existence of ground truth images enables an objective evaluation while the provided image sequences are useful to assess the temporal consistency of the recovered images. Our preliminary results reveal that further developments and improvements of dehazing algorithms are required to recover natural colors and to meet a professional image quality level.

References

[1] Tan, R.: Visibility in bad weather from a single image. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008. (June 2008) 1–8

[2] Fattal, R.: Single image dehazing. ACM Trans. Graph. **27**(3) (August 2008) 72:1–72:9

[3] He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**(12) (Dec 2011) 2341–2353

[4] Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: IEEE International Conference on Computer Vision (ICCV), 2013. (Dec 2013) 617–624

[5] Narasimhan, S., Nayar, S.: Chromatic framework for vision in bad weather. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2000. Volume 1. (2000) 598–605

[6] Nayar, S., Narasimhan, S.: Vision in bad weather. In: IEEE International Conference on Computer Vision (ICCV), 1999. Volume 2. (1999) 820–827

[7] Tang, K., Yang, J., Wang, J.: Investigating haze-relevant features in a learning framework for image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014. (June 2014) 2995–3002

[8] Ancuti, C., Ancuti, C.: Single image dehazing by multi-scale fusion. IEEE Transactions on Image Processing **22**(8) (Aug 2013) 3271–3282

[9] Fattal, R.: Dehazing using color-lines. ACM Trans. Graph. **34**(1) (December 2014) 13:1–13:14

[10] Nishino, K., Kratz, L., Lombardi, S.: Bayesian defogging.

SSIM	0.435
VIF	0.189
FSIMc	0.601
CID	0.817
e	-
$\bar{r}$	-

SSIM	0.436
VIF	0.086
FSIMc	0.549
CID	0.850
e	-
$\bar{r}$	-

SSIM	0.496
VIF	0.203
FSIMc	0.682
CID	0.803
e	2.326
$\bar{r}$	3.560

SSIM	0.421
VIF	0.240
FSIMc	0.657
CID	0.759
e	374.719
$\bar{r}$	8.169

SSIM	0.614
VIF	0.259
FSIMc	0.741
CID	0.663
e	1.849
$\bar{r}$	2.758

SSIM	0.597
VIF	0.401
FSIMc	0.703
CID	0.707
e	308.582
$\bar{r}$	5.446

SSIM	0.537
VIF	0.246
FSIMc	0.703
CID	0.728
e	1.625
$\bar{r}$	2.421

SSIM	0.473
VIF	0.139
FSIMc	0.557
CID	0.836
e	36.087
$\bar{r}$	2.595

SSIM	0.595
VIF	0.265
FSIMc	0.725
CID	0.650
e	1.604
$\bar{r}$	2.782

SSIM	0.590
VIF	0.356
FSIMc	0.684
CID	0.726
e	172.597
$\bar{r}$	4.561

SSIM	0.654
VIF	0.499
FSIMc	0.768
CID	0.790
e	285.093
$\bar{r}$	7.781

Figure 8: Restoration results and quality metrics for the *Greenhouse* and *Innolab* images. Top row: Ground truth color image. Second row: Hazy input image. Third row: Result of Nishino et al. [10]. Fourth row: Result of Meng et al. [4]. Fifth row: Result of Tarel & Hautière [11]. Sixth row: Result of Tang et al. [7]. Last row: Result of Connah et al. [17] (where available).

Int. J. Comput. Vision **98**(3) (July 2012) 263–278
[11] Tarel, J.P., Hautire, N.: Fast visibility restoration from a single color or gray level image. In: Computer Vision, 2009 IEEE 12th International Conference on. (Sept 2009) 2201–2208
[12] Kratz, L., Nishino, K.: Factorizing scene albedo and depth from a single foggy image. In: Computer Vision, 2009 IEEE 12th International Conference on. (Sept 2009) 1701–1708
[13] Feng, C., Zhuo, S., Zhang, X., Shen, L., Susstrunk, S.: Near-infrared guided color image dehazing. In: IEEE International Conference on Image Processing (ICIP), 2013. (Sept 2013) 2363–2367
[14] Schechner, Y., Narasimhan, S., Nayar, S.: Instant dehazing of images using polarization. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2001. Volume 1. (2001) 325–332
[15] Schaul, L., Fredembach, C., Susstrunk, S.: Color image dehazing using the near-infrared. In: IEEE International Conference on Image Processing (ICIP), 2009. (Nov 2009) 1629–1632
[16] Zhang, B., Wang, L.: Local color transfer based on dark-channel dehazing for visible/infrared image fusion. In: Proc. SPIE 8056, Visual Information Processing XX. (2011)
[17] Connah, D., Drew, M., Finlayson, G.: Spectral edge image fusion: Theory and applications. In: European Conference on Computer Vision (ECCV), 2014. Volume 8693 of Lecture Notes in Computer Science. Springer International Publishing (2014) 65–80
[18] McCartney, E.J.: Optics of the Atmosphere. Wiley and Sons, Inc (1976)
[19] Narasimhan, S., Nayar, S.: Vision and the atmosphere. International Journal of Computer Vision (IJCV), 2002 **48**(3) (2002) 233–254
[20] Narasimhan, S.G., Wang, C., Nayar, S.K.: All the images of an outdoor scene. In: Proceedings of the 7th European Conference on Computer Vision-Part III, London, UK, Springer-Verlag (2002) 148–162
[21] Brown, M., Susstrunk, S.: Multi-spectral sift for scene category recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011. (June 2011) 177–184
[22] Tarel, J.P., Hautiere, N., Cord, A., Gruyer, D., Halmaoui, H.: Improved visibility of road scene images under heterogeneous fog. In: IEEE Intelligent Vehicles Symposium (IV), 2010. (June 2010) 478–485
[23] ITU: Recommendation ITU-R BT.709-5: Parameter values for the HDTV standards for production and international programme exchange (Apr 2002)
[24] Zilly, F., Eisert, P., Kauff, P.: Real-time analysis and correction of stereoscopic HDTV sequences. Proceedings of Visual Media Production (2009)
[25] Hautière, N., Tarel, J.P., Aubert, D., Dumont, E.: Blind contrast enhancement assessment by gradient ratioing at visible edges. Image Analysis & Stereology Journal **27**(2) (June 2008) 87–95
[26] Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4) (April 2004) 600–612
[27] Sheikh, H.R., Bovik, A.C.: Image information and visual quality. IEEE Transactions on Image Processing **15**(2) (Feb 2006) 430–444
[28] Zhang, L., Zhang, L., Mou, X., Zhang, D.: Fsim: A feature similarity index for image quality assessment. IEEE Transactions on Image Processing **20**(8) (Aug 2011) 2378–2386
[29] Preiss, J., Fernandes, F., Urban, P.: Color-image quality assessment: From prediction to optimization. IEEE Transactions on Image Processing **23**(3) (March 2014) 1366–1378

Author Biography

Julia Lüthen received her M.Sc. degree in electrical and computational engineering from the Technical University of Munich in 2015. Since 2013 she works for ARRI Cine Technik as Image Processing Engineer in Research and Development.

Julian Wörmann received his M.Sc. degree in electrical and computational engineering from the Technical University of Munich in 2012, where he is currently pursuing the doctoral degree at the Department of Electrical and Computer Engineering. His current research interests include sparse signal models, geometric optimization, inverse problems in image processing, and computer vision.

Martin Kleinsteuber received his PhD in Mathematics from the University of Würzburg, Germany, in 2006. After post-doc positions at National ICT Australia Ltd, the Australian National University, Canberra, Australia, and the University of Würzburg, he joined TU München, Germany, in 2009 as assistant professor for geometric optimization and machine learning at the Department of Electrical and Computer Engineering. In 2016, he has been appointed Lead Data Scientist at Mercateo AG, Munich.
He won the SIAM student paper prize in 2004 and the Robert-Sauer-Award of the Bavarian Academy of Science in 2008 for his works on Jacobi-type methods on Lie algebras. His research interests are in the areas of statistical signal processing, machine learning and computer vision.

Johannes Steurer is Principal Engineer R&D at ARRI Cine Technik and responsible for research and innovations such as depth-sensing, self-localization, lens focusing for motion picture cameras. Steurer received his Dr.-Ing. degree in Electrical and Computer Engineering from Technical University Munich in 1992. Scientific lecturer and expert reviewer for various institutions, member of SMPTE and FKTG. He received several awards including an OSCAR®Statuette for the development of the ARRILASER film recorder.