

Optimizing Postprocessing of Range-Gated Viewing Data for Maritime Search and Rescue Operations at Night and in Bad Weather Conditions

Tristan Preis¹, Jendrik Schmidt¹, Alexander Klein¹, Enno Peters¹, Thomas Lübcke², Maurice Stephan¹

At night and in bad weather conditions the detection of persons and objects floating in the sea represents a major challenge for search and rescue operations (SAR). If conventional searchlights are used, backscattering from rain, fog and snow decreases detection range. Therefore, a compact and inexpensive range-gated viewing system which significantly reduces atmospheric backscattering was developed. The instrument was designed for detection ranges of several hundred meters. In this study, different image processing techniques were analyzed in terms of improved object detectability for a human observer and for a machine learning-based object detector, based on a real-world image dataset. On the one hand, noise of the camera is reduced by performing a non-uniformity correction (NUC) and on the other, the dynamic range of the images is adjusted and dark objects are accentuated by equalizing (EQ). The aim of this field study with the subsequent post processing steps was to improve visibility for both human observers and machine learning-based object detectors with low computing power, based on real-world image datasets. The results show that processing requirements are different in both cases, mainly due to human eye perception, which an automated detector does not rely on and therefore the performance of the object detector before the equalizing step is slightly better. However, the NUC improves the image quality in any case.

KEY WORDS

- ~ Vision enhancement
- ~ Non-uniformity correction (NUC)
- ~ Image processing
- ~ Gated-viewing
- ~ Range-gated viewing
- ~ Equalization
- ~ Object detection (YOLOv8)
- ~ Maritime search and rescue operation (SAR)

¹ German Aerospace Center, Institute for the Protection of Maritime Infrastructures, Bremerhaven, Germany.

² German Maritime Search and Rescue Service, Bremen, Germany

e-mail: Tristan.Preis@dlr.de

doi: 10.7225/toms.v13.n01.w17

Received: 22 Dec 2023 / Revised: 22 Feb 2024 / Accepted: 7 Mar 2024 / Published: 15 Mar 2024

This work is licensed under



1. INTRODUCTION

Even with a normal searchlight in the visible spectrum, objects with retroreflectors can be spotted at large distances far better than without reflectors. For this reason, many life vests and other objects designed to be seen, are equipped with retroreflectors (International Maritime Organization, 1989). However, in some cases, persons overboard do not wear special clothing with retroreflectors, causing a dimmer reflection in the image and making their detection more difficult, especially in harsh weather conditions. As a result, the main objective of this study is to improve the visibility for operators participating in search and rescue operations (SAR) at night and in scattering environments (fog, rain, snow) through optical instruments. In cooperation with the German Maritime Search and Rescue Service (Deutsche Gesellschaft zur Rettung Schiffbrüchiger, DGzRS) a field test comparing different camera technologies was performed in the German North Sea in the vicinity of Bremerhaven, onboard the DGzRS rescue cruiser ‘Hermann Rudolf Meyer’, on December 6, 2022 during nighttime (darkness).

As available space and load on a rescue cruiser are limited, system size and weight should be as small as possible. In addition, many rescue services depend on donations, and therefore require cost efficient solutions and instruments. Therefore, a Transportable Range-Gated Viewing System (TRAGVIS)¹ was developed to detect persons overboard and other floating objects (Peters et al., 2019). In addition, the device should be simple to use and images easy to interpret. For comparison purposes, a thermal (IR) camera and a passive, monochrome visible-range (VIS) camera were used in addition to the range-gated viewing system.

Different scenes were recorded containing a dummy without retroreflectors and the rescue vessel’s daughter boat. An example of a scene from the field test is shown in Figure 1. The figure shows identical scenes recorded simultaneously by the range-gated viewing system (Fig. 1a), the thermal camera (Fig. 1b) and the passive visible light camera (Fig. 1c), respectively. All images are raw data without any post-processing applied.

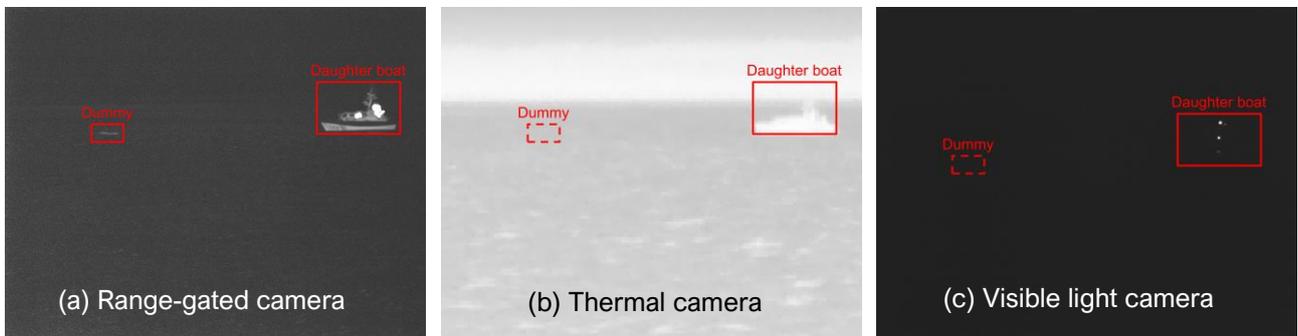


Figure 1. The same scene observed with (a) the range-gated viewing system, (b) thermal camera and (c) visible light camera. All images are raw data, i.e. without post-processing

The range-gated image in Fig. 1(a) shows the rescue vessel’s daughter boat and the dummy without retroreflectors. The majority of the image appears to be dark, as most of the light emitted is reflected from the water surface into the sky. In the thermal image (Fig. 1(b)), the daughter boat can be clearly identified, but the dummy is not distinguishable from the surrounding waves. While the thermal imaging camera is hardly affected by the backscatter, the image only shows the thermal distributions and not the objects, which is crucial when it comes to recognizing objects that have been floating in the sea for some time. In the visible image (Fig. 1(c)) taken without active illumination, everything is dark except for the navigational lights of the daughter boat. In comparison to the image recorded with TRAGVIS (Fig. 1(a)), the images shown in Figure 1(b) and Figure 1(c)

¹ TRAGVIS hardware is described in more detail in Section 2.

are useless for search and rescue operators as they do not help detect the dummy (person overboard). However, even in the gated image, the dummy is difficult to detect, especially when considering that search and rescue operations are often not performed in calm (as in this test) but in rough seas. This is why this study aims to further improve TRAGVIS images, and allow even better (and more reliable) detection of persons and objects overboard by human operators and automated object detectors.

The study is structured as follows: range-gated viewing instrument and measurement technique are briefly explained in Section 2. Post-processing techniques improving the inspection of images are presented in Section 3, focusing on better recognition by human operators. In addition, an automated, computer-assisted object detector is introduced in Section 4, tested to support human search and rescue operators. The algorithm was trained using a labeled dataset from the field test. After training, a selection of test images was used to quantify the performance of the object detector. In Section 5, processing step results are evaluated and discussed with respect to different object observers (human and automated detector). Section 6 gives conclusion and outlook.

2. INSTRUMENT AND MEASUREMENT TECHNIQUE

The measuring setup and technique of range-gated viewing are shown in Figure 2. Fig. 2(a) is a photo of the instrumental setup used on the rescue cruiser in the field test. The system consists of two housings: one for the visible (Onyx from Teledyne e2v) and thermal (VarioCAM HDx from InfraTec) camera, which are both passive cameras, and one for the range-gated viewing instrument, which consists of a special camera (Bora 2D from Teledyne e2v) equipped with a fast electronic shutter and a self-developed pulsed laser illuminator. The entire system was placed on a pan-tilt-unit (PTU) for sea stabilization and remote control.

The range-gating (also known as gated-viewing) technology attempts to physically reduce backscattering from rain, fog or snow in front of the sensor system (Laurenzis, 2012). The measuring technique involves illuminating and recording a scene at a specific depth level, as illustrated in Fig. 2(b). A pulsed laser illuminator is synchronized with a camera, that can open and close its electronic shutter typically within nanoseconds. The adjustable delay between laser pulse emission and camera shutter opening eliminates non-essential parts of an image, which then consists of a defined gate, that is exposed.

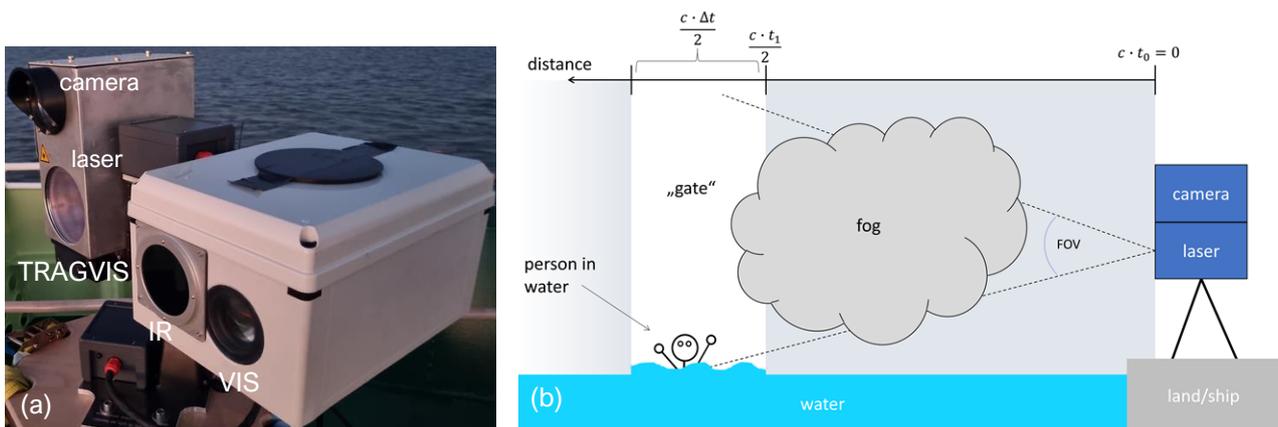


Figure 2. (a) Measuring setup on the rescue cruiser with range-gated viewing system consisting of camera (top) and laser illuminator (bottom) (TRAGVIS), thermal and VIS camera; (b) Range-gated viewing measuring technique.

Following laser pulse emission at time t_0 , the camera shutter stays closed, while the pulse propagates towards the gate. Due to the closed shutter, backscattering from particles along the light path is not recorded

by the camera, i.e. the area between the instrument and the gate stays dark in the recorded image. At time t_1 the camera shutter opens and the first reflections of the pulse are recorded. When the camera and laser are at the same location, the distance of the rising edge of the gate d_{gf} is half of time t_1 multiplied by the speed of light c :

$$d_{gf} = c \cdot t_1/2 \dots \dots \dots (1)$$

Similarly, gate width Δt is defined by the width of the laser pulse and of the open camera shutter. This process is repeated ca. 17,000 times (approx. 50 ms) to accumulate enough photons before reading the camera (which again only takes milliseconds). The effective framerate the instrument provides is therefore 16-17 images/sec.

Many range-gated viewing systems operate in the short-wavelength infrared (SWIR) spectrum at 1.57 μm and new approaches tend to use the 2.09 μm spectrum (Göhler et al., 2023). The advantage of these spectra is the reduction of atmospheric scintillation, however, they are associated with high costs as sensors are made from indium gallium arsenide (InGaAs) or mercury cadmium telluride (MCT). TRAGVIS operates at approximately 808 nm and is equipped with a CMOS camera, which has a sensor with a good quantum efficiency of 42 % in the near-infrared (NIR) spectrum. This design decision was made to reduce construction costs and because Mie calculations have shown that longer wavelengths (SWIR) have no advantage in terms of scattering and extinction efficiencies when fog particles consisting of water are large (Peters et al., 2023), which is typically the case in the environments TRAGVIS was designed for. Interestingly, Bett et al. (2023) proposed using a single-pixel camera, which could potentially reduce costs. However, given that the image would then have to be generated with the help of a neural network, which requires computing power and training datasets, the traditional approach was used in this study.

As the main objective of TRAGVIS is to find people in distress at sea, a wide field of view (FOV) of 7° x 6° was chosen (which roughly matches the FOV of standard binoculars) to cover a large area in each image. The maximum detection range in fog conditions is approximately 250 m. To address eye-safety as a major concern, a self-built beam expander was developed. The resulting instrument meets laser class 1M system requirements.²

Previous measurements have shown that in moderate aerosol conditions that prevailed during the field test, the first 50m contribute the most to image backscattering. Consequently, the images recorded in this paper suppress the first 50m only, which is sufficient for an improved image and enabling wide gate due to the speed of full scene recording. Normally, a range-gating instrument provides a sequence of images, in which the position of the gate varies and images need to be combined. This enables obtaining depth information, but depth information is not needed for the purpose of this study.

3. POST-PROCESSING

Gated-viewing images usually suffer from large fixed pattern noise that decreases image quality. After acquiring an image, a non-uniformity correction (NUC) is performed. A corrected 1-point NUC x_{1-NUC} image is calculated by subtracting a reference dark image $x_{ref,dark}$ from the recorded image x_{obj} , which compensates for the so-called dark signal non-uniformity (DSNU):

$$x_{1-NUC} = x_{obj} - x_{ref,dark} \dots \dots \dots (2)$$

The dark image is acquired with a mechanically closed instrument preventing any light from reaching the sensor. In addition, the dark image is the average of 100 single exposures, intended to average out temporal noise, so that the final dark image consists of spatial patterns only.

² Further technical information can be found in Peters et al., 2019

After the subtraction of the dark image, the mean gray value $\overline{x_{ref,dark}}$ can be added to all pixels resulting in an image $x_{1-NUC,mean}$ having the same background gray value:

$$x_{1-NUC,mean} = x_{obj} - x_{ref,dark} + \overline{x_{ref,dark}} \dots \dots \dots (3)$$

The idea of a further 2-point NUC is to reduce spatial noise, caused by the different sensitivities of the pixels when exposed to light, which leads to the so-called Photo-Response Non-Uniformity (PRNU). Therefore, matrix m is introduced. The matrix is obtained from a reference image $x_{ref,mid}$, in which the sensor is illuminated by a homogenous light (usually at approx. 50% saturation), and a reference dark image $x_{ref,dark}$:

$$m = (\overline{x_{ref,mid}} - \overline{x_{ref,dark}}) / (x_{ref,mid} - x_{ref,dark}) \dots \dots \dots (4)$$

After multiplying an uncorrected raw image, from which the dark image has been subtracted, with m , all pixel photo responsivities have the same slope³. As for the 1-point NUC, the mean gray value of the dark image might ($x_{2-NUC,mean}$) or might not (x_{2-NUC}) be added to the corrected image:

$$x_{2-NUC,mean} = (x_{obj} - x_{ref,dark})m + \overline{x_{ref,dark}} \dots \dots \dots (5)$$

$$x_{2-NUC} = (x_{obj} - x_{ref,dark})m \dots \dots \dots (6)$$

To highlight more details, the image can be equalized (EQ). Most parts of images recorded during night over water are usually dark. Since simply brightening the scene results in a milky appearance useless to the human operator, a more sophisticated solution is required. In addition, as the instrument gives 12-bit raw data, and the displays are 8-bit, thus a conversion from 12 bit to 8 bit is required. As diffused reflections from distant objects are usually very small, low signals should be brightened up more than large signals in order to increase visibility for human operators. Therefore, an equalizer algorithm is adopted, that is usually used for astrophotography⁴. The equalization process can be separated into two functions $f(x)$ and $g(x)$, which are called one after the other.

The function $f(x)$ boosts dark gray values towards brighter values, as shown in Figure 3. Details on the dark gray scale have higher contrast. Function $g(x)$ pushes dark and bright pixel values to the middle of the gray value range.

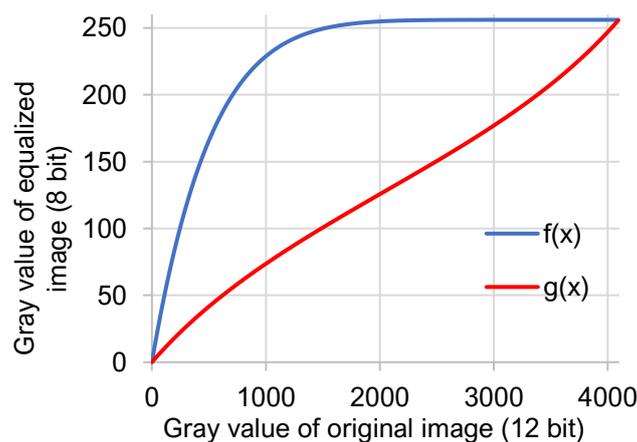


Figure 3. Equalization graphs including 12 bit to 8 bit conversion

³ The 2-point NUC suggested here assumes a linear photo-response.

⁴ <https://yager.io/Astro.html#Post-processing>, 22.02.2024

Table 1 summarizes all combinations of corrections and equalizations performed during post-processing. To identify the optimal method, all processing combinations were performed for each image recorded by TRAGVIS, resulting in ten different images of the same scene, which are then compared to each other.

Processing configuration
Original – raw TRAGVIS image
1-point NUC
2-point NUC
1-point NUC + mean
2-point NUC + mean
Original + EQ
1-point NUC + EQ
2-point NUC + EQ
1-point NUC + mean + EQ
2-point NUC + mean + EQ

Table 1. List of processing steps and resulting images 1. List of processing steps and resulting images

4. AUTOMATED OBJECT DETECTION

Apart from manual detection by human observers, automated, machine learning-based object detectors can assist the SAR forces. The image processing requirements for automated detectors may differ from what a human observer defines as good image quality. Therefore, the same set of images with processing configurations was tested using a machine learning-based object detector.

A large number of different object detection algorithms is available nowadays. Models such as Region-based Convolutional Neural Networks (R-CNN) (Girshick, Ross, et al., 2014) use a two-stage process, where general object areas are first proposed, followed by a second set where the object is classified and localization is further refined. Currently, detection transformer models give the best results (Carion et al., 2020). Given the need to apply the detector on limited hardware in real time, SAR purposes require the detector to be as fast as possible, while delivering accurate results. As a detailed evaluation on the recorded, relatively small dataset is not meaningful between different object detectors, this paper focuses on the use of the latest version of the YOLO object detector, specifically YOLOv8 (Ultralytics, 2023) for automatic detection tasks, which is a widely used and well-known algorithm. YOLO uses only a single end-to-end pipeline to establish object class and bounding box, which makes it one of the fastest modern object detectors.

This paper compares YOLOv8n with 3.2 million and YOLOv8x with 68.2 million parameters. Both models were pretrained on the COCO dataset, which is integrated in the YOLO algorithm. Pretraining is used to set the object detector in advance so that it can then recognize edges and other geometric structures. As pretraining does not aim to detect ships and other maritime objects, it does not matter whether the object detector analyzes images at night or during the day. After pretraining, the training dataset is used to fully train the object detector for 100 epochs using the full image size of 1280 x 1024 pixels. For training and evaluation, the dataset was split in 461 training and 38 test images. The training and test datasets were also expanded with each processing step (s. Tab. 1), so that in the end 4,610 images for training and 380 images for evaluation are available. All the training data were learned together in the object detector. The test set was split up for each processing configuration given in Table 1, to compare their effects on detection quality.

5. RESULTS

As mentioned in Section 4, 499 different scenes were recorded with TRAGVIS in the field test. The scene already shown in Fig. 1 is used to demonstrate the effect of the processing techniques described in Section 3. Fig. 4 summarizes different combinations using a 1-point NUC. Due to the subtraction of the reference dark image $x_{ref,dark}$, the resulting image in Fig. 4a for the 1-point NUC is generally darker. For human observers, the recognition of the dummy against the background gets worse, as the human eye distinguishes contrasts between dark gray levels less efficiently compared to bright gray levels on most computer displays and monitors (Kimpe and Tuytschaever, 2007). Therefore, the dummy is more easily identified when the mean gray value of the reference dark image is added as in Fig. 4b. The results after equalization are shown in Fig. 4c and Fig. 4d. A significant improvement can be observed in the image with 1-point NUC and equalization without the addition of the average (Fig. 4d), where dummy identification is easier compared to the other processing configurations.

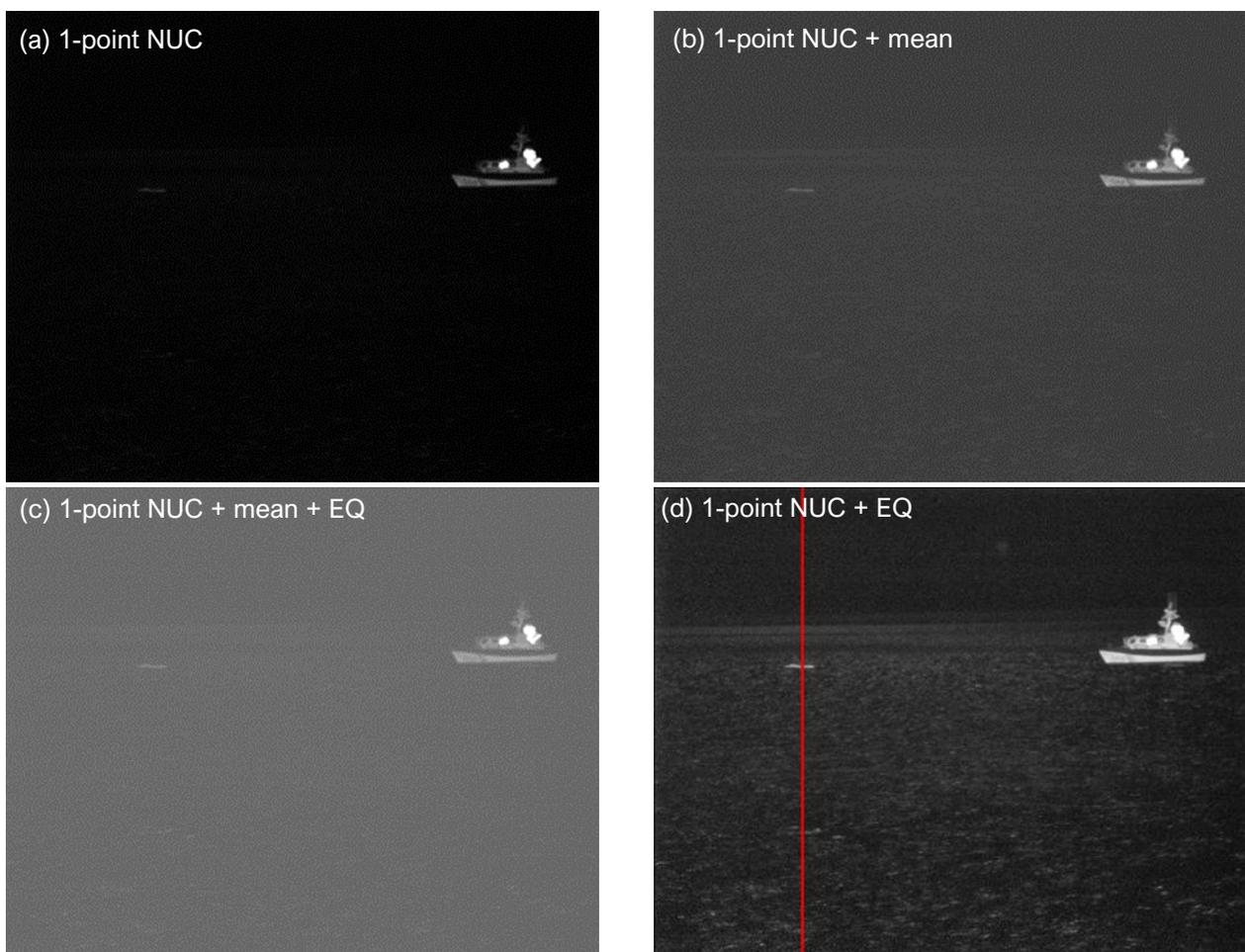


Figure 4. Range-gated images with different preprocessing configurations: (a) 1-point NUC, (b) 1-point NUC with added mean gray value (c) 1-point NUC with added mean gray value after equalization and (d) 1-point NUC after equalization with vertical cross section to increase the contrast of the dummy against the background (s. Fig. 5).

To quantify the effect of different preprocessing configurations on dummy detection, the gray values along a vertical line passing through the position of the dummy were analyzed. In Figure 5, the cross-section along the line indicated in Fig. 4d is shown. The dummy is visible with a gray value of 148, compared to the

average of 42 along the indicated line. The Michelson contrast (C_M) (Michelson, 1927) was calculated based on the dummy gray value being higher than the background gray value. The Michelson contrast was calculated between the dummy (x_{max}) and the average of the background (x_{med}):

$$C_M = (x_{max} - x_{med}) / (x_{max} + x_{med}) \dots \dots \dots (6)$$

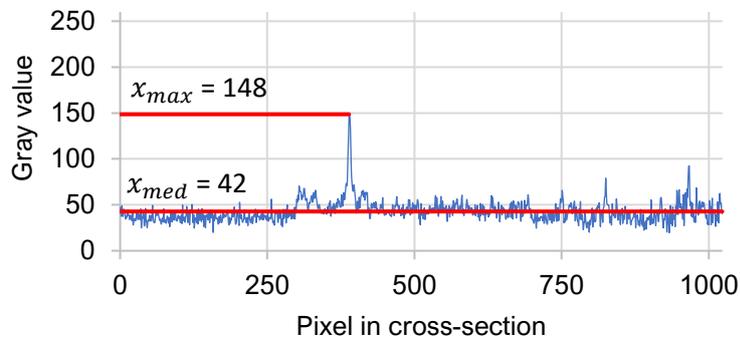


Figure 5. Gray values along the red line in Fig. 4d

The contrast in the original image and the images obtained after the non-uniformity correction with the added mean value is similar. Not adding the mean gray value in the NUC increases the contrast significantly. However, this has to be taken with reservation because, as previously mentioned, the human eye better perceives differences between bright gray levels on computer displays. That can be seen in Fig. 4a (1-point NUC) and b (1-point NUC + mean): the dummy is almost invisible to the human observer in Fig. 4a, while it is much easier to identify when the background mean is added as in Fig. 4b. However, the calculated Michelson contrast in Table 2 is 0.86 for Fig. 4a and is as low as in the original image, with the value of 0.23 for the configuration shown in Fig 4b. The reason lies in the definition of the Michelson contrast (eq. 6). The lower the background gray value, the closer to 1 the contrast will get. To conclude, the Michelson contrast is not a good postprocessing measure in case of a human observer.

Processing step	Michelson contrast	Precision	Recall
Original	0.22	0.94	0.83
1-point NUC	0.86	0.96	0.83
2-point NUC	0.85	0.96	0.86
1-point NUC + mean	0.23	0.96	0.82
2-point NUC + mean	0.22	0.96	0.83
Original + EQ	0.13	1.00	0.75
1-point NUC + EQ	0.71	0.95	0.76
2-point NUC + EQ	0.71	0.93	0.85
1-point NUC + mean + EQ	0.12	0.92	0.83
2-point NUC + mean + EQ	0.12	0.88	0.90

Table 2. Michelson contrast for each processing step and object detection results (precision and recall) of test images for YOLOv8n for class ‘Fixed and Floating Objects’ (i.e. dummy)

The 2-point NUC usually improves image quality more than the 1-point NUC. In this study however, the 2-point NUC and the 1-point NUC gave very similar results. The reason is that the images are mostly dark and

the influence of pixel sensitivity is negligible compared to the influence of the dark signal non-uniformity. By contrast to a human observer, a machine learning-based detector performs best using a different configuration of postprocessing steps.

Table 2 gives an overview of the precision and recall for the dummy when using the object detector. Precision is defined as correctly detected objects divided by the total number of detections by the object detector (including false detections). In contrast, recall is the share of correctly detected objects in all the relevant objects. Relevant objects are the number of labeled objects in the test dataset.

A comparison of object detector results with the Michelson contrast in Table 2 suggests that the object detector tends to perform better when the Michelson contrast is high, most likely due to better edge detection. The best results are obtained by the 1-point NUC and 2-point NUC before equalization, as they have the highest contrast. All other processing variants are tradeoffs between precision and recall compared to the original raw image.

6. CONCLUSION

The low-cost and lightweight range-gated viewing system used in the study has an advantage in finding persons in distress and other floating objects in the sea over thermal imaging and visible light records. In this study, non-uniformity correction and equalizer algorithm were used, two image post-processing technologies that require low computing power and can therefore also be run on small computing units. Range-gated viewing (on the hardware side) and non-uniformity correction and equalization (on the software side) improve object detection for both human search and rescue operators and automated object detectors. Nevertheless, different postprocessing configurations were found to give best results.

In case of a human operator, 1-point NUC and 2-point NUC after equalization are the best option, because shifting the dark dummy to the middle of the grayscale supports human perception. In addition, the Michelson contrast was found not to be a good measure for identifying the optimal postprocessing configuration for human perception, which is specific to situations involving dark objects of interest against the dark background. In contrast, for automated object detection, mean gray level of the image does not play a role in detection capability. Processing steps, which produce images with the highest contrast, deliver the best results for automated object detection. In this instance, the 1-point NUC and 2-point NUC before equalization deliver the best results, given that equalization, although improving human perception, decreases the Michelson contrast.

The 2-point NUC did not perform significantly better than the 1-point NUC in none of the two object detection methods (given that the image is mainly dark and thus suffering predominantly from DSNU), but is more complex and requires higher computational power. Therefore the recommendation is to perform the 1-point NUC only, especially when space and computing power are an issue, as in systems used on rescue vessels.

As an outlook, future research should focus on other equalizer algorithms. Apart from the techniques used, other mechanisms, like filtering and masking for vision enhancement, should also be evaluated.

ACKNOWLEDGMENTS

We acknowledge that the crew of the search and rescue vessel 'Hermann Rudolf Meyer' from the German Maritime Search and Rescue Service (DGzRS) supported the field test. The technical and scientific work presented here, was funded by the Federal Ministry for Economic Affairs and Climate Action of Germany (Bundesministerium für Wirtschaft und Klimaschutz, BMWK) and the state of Bremen.

REFERENCES

- Bett, C. et al., 2023. Evaluation of a time-gated-single-pixel-camera as a promising sensor for autonomous vehicles in harsh weather conditions. *J. Eur. Opt. Society-Rapid Publ.* 19(1), p.27. Available at: <https://dx.doi.org/10.1051/jeos/2023023>
- Carion, N. et al., 2020. End-to-end object detection with transformers. In *European conference on computer vision*, pp. 213-229. Cham: Springer International Publishing.
- Girshick, R. et al., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580-587
- Göhler, B. et al., 2023. Gated viewing at 2.09 μm laser wavelength: experimental system assessment and comparison to 1.57 μm , *Proc. SPIE 12737, Electro-Optical and Infrared Systems: Technology and Applications XX*. Available at: <https://doi.org/10.1117/12.2684727>
- International Maritime Organization, 1989. RESOLUTION A.658(16) - Use and Fitting of Retro-Reflective Materials on Life-Saving Appliances. Available at: [https://wwwcdn.imo.org/localresources/en/KnowledgeCentre/IndexofIMOResolutions/AssemblyDocuments/A.658\(16\).pdf](https://wwwcdn.imo.org/localresources/en/KnowledgeCentre/IndexofIMOResolutions/AssemblyDocuments/A.658(16).pdf)
- Kimpe, T., and Tuytschaever, T., 2007. Increasing the number of gray shades in medical display systems—how much is enough? *J. Digit. Imaging* 20(4), pp. 422-32. Available at: <https://dx.doi.org/10.1007/s10278-006-1052-3>
- Laurenzis, M., 2012. Investigation of range-gated imaging in scattering environments. *Optical Engineering* 51. Available at: <https://dx.doi.org/10.1117/1.OE.51.6.061303.5>
- Michelson, A.A., 1927. *Studies in optics*. University of Chicago Press.
- Peli E., 1990. Contrast in complex images. *J. Opt. Soc. Am. A* 7(10), pp. 2032-40. doi: Available at: <https://dx.doi.org/10.1364/josaa.7.002032>
- Perry, D. and Dereniak, E., 1993. Linear theory of nonuniformity correction in infrared staring sensors. *Optical Engineering*, 32, pp. 1854-1859. Available at: <https://dx.doi.org/10.1117/12.145601>
- Peters, E., et al., 2023. Performance studies of two active SWIR instrument designs for vision enhancement in indoor applications, *Proc. SPIE 12737, Electro-Optical and Infrared Systems: Technology and Applications XX*, Available at: <https://dx.doi.org/10.1117/12.2680281>
- Peters, E., et al. 2019. Development of a novel low-cost NIR gated-viewing sensor for maritime search and rescue applications, 3. Available at: <https://dx.doi.org/10.1117/12.2532538>
- Ultralytix, 2023. Available at: <https://docs.ultralytix.com>