Vision, Modeling, and Visualization (2023)
T. Grosch and M. Guthe (Eds.)

Optimizing Temporal Stability in Underwater Video Tone Mapping

M. Franz, B.M. Thang, P.Sackhoff, T. Scholz, J. Möller, S. Grogorick and M. Eisemann

TU Braunschweig, Germany

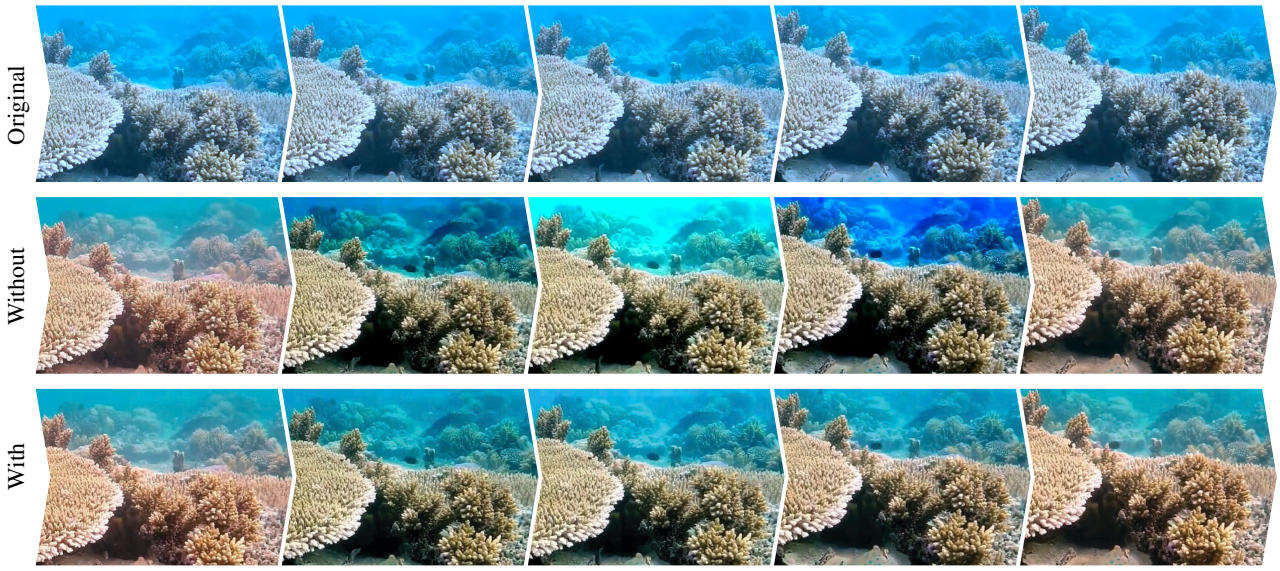

Figure 1: *Example sequence (top), tone-mapped without (middle) and with our temporal stabilization extension (bottom).*

Abstract

In this paper, we present an approach for temporal stabilization of depth-based underwater image tone mapping methods for application to monocular RGB video. Typically, the goal is to improve the colors of focused objects, while leaving more distant regions nearly unchanged, to preserve the underwater look-and-feel of the overall image. To do this, many methods rely on estimated depth to control the recolorization process, i.e., to enhance colors (reduce blue tint) only for objects close to the camera. However, while single-view depth estimation is usually consistent within a frame, it often suffers from inconsistencies across sequential frames, resulting in color fluctuations during tone mapping. We propose a simple yet effective inter-frame stabilization of the computed depth maps to achieve stable tone mapping results. The evaluation of eight test sequences shows the effectiveness in a wide range of underwater scenarios.

CCS Concepts

• ***Computing methodologies*** → *Perception; Image processing;*

1. Introduction

Imagery of underwater scenes often appears blurry and noisy due to light scattering from turbidities in the water, and suffers from strong color shift that occurs in underwater images, because colors are attenuated depending on their wavelength. Long wavelength light (red) is already absorbed at a water depth or distance of about 3 m. Light of shorter wavelengths (green and blue) is only absorbed at greater distances or deeper depths, yielding the typical blue-green color cast known from most underwater images. There are many works in the literature that improve the visual quality of such underwater scenes. Part of these approaches is based on the evaluation of depth information, to improve the colors of focused objects, while

DOI: 10.2312/vmv.20231239

maintaining the blue tint for more distant regions. The drawback of such approaches, however, is the heavy dependency on valid and temporally stable depth information. While approaches for single frame depth estimation are continuously improving, and can be assumed to be sufficiently consistent for such a guiding task, temporal stability is often an issue. Many such techniques suffer from inconsistencies across sequential frames, e.g., because the employed depth estimation is based on image sharpness, which can easily be distorted by motion blur. This can cause the tone mapping to produce unwanted color fluctuations when processing video.

To achieve results without visible flicker in generated videos, we propose an extension for existing single frame recoloring methods, that rely on depth estimation to enhance underwater images [PZC15; PC17]. The basic idea is to optimize depth maps of individual frames according to the corresponding optical flow of preceding and subsequent frames. Using the optical flow information, i.e., how pixel positions move across frames, we can keep the computed depth maps temporally coherent with previous frames. We achieve improvements both for individual frames, which obtain more accurate depth information, as well as for the entire image sequence, by diminishing disturbing temporal artifacts.

Using 8 different video sequences showing various underwater scenarios, we test our approach with two single-frame methods. We test both methods as they are, and extended with our stabilization approach. We assess the results according to Panetta et al. [PGA15], using established metrics for underwater image quality (UISM and UICM) via Krahn et al. 's evaluation framework [KM16].

The results show that our approach noticeably reduces inconsistencies in color between consecutive frames. Furthermore, it obtains comparable or better results with respect to colorfulness and image sharpness.

In the following, we will first give an overview of the research field in section 2, followed by a description our approach, and how it is used to extended single-frame methods for more stable results on videos in section 3. Finally, we compare two exemplary single-frame methods with and without our multi-frame temporal stabilization extension, with respect to their effect on underwater image quality in section 4.

2. Related Work

The problem of underwater video enhancement has been addressed in several previous works, which can roughly be divided into two classes: prior-based and learning-based methods. The latter have gained a lot attention in recent years, also in the field of underwater image enhancement, mostly in the form of convolutional neural networks (CNNs) and generative adversarial networks (GANs) [ZYZ23; HWZ*22].

Regarding learning-based methods, UIE-Net by Wang et al. [WZCW17] is an example of CNN-based approaches, that consists of two subnets dedicated to color correction and haze removal, respectively. Similarly, Fu and Cao [FC20] also combine two subnets for color and contrast optimization, but compliment this with a classical hand-crafted histogram equalization at the end of their pipeline. The UWCNN by Li et al. [LAP20] is a single model trained to directly improve the quality of underwater footage in just one model. Lu et al. [LLU*18] present a specialized CNN for degraded low-light underwater images.

Approaches based on GANs include the UGAN by Fabri et al. [FIS18], that employs the approach of the CycleGAN by Thu et al. [ZPIE17] to generate training pairs instead of manually creating training data. The WaterGAN by Li et al. [LSEJ17] generates underwater images from in-air RGB-D images, to then train an appropriate correction network in a second step. Islam et al. [IXS20] generate a dataset of more than 30000 images, that is used to train a model, dedicated to improve underwater object detection and human body-pose estimation performance.

Many prior-based methods are based on the so-called dark channel prior (DCP) for haze-reduction, assuming that, typically, each pixel (except for the sky) has at least one color channel with very low intensity. Chiang et al. [CC12] exploit that to estimate the distance of objects to the camera, to compensate for respective color distortions. A specialized version, called underwater DCP (UDCP), is proposed by Drews et al. [DNM*13] to specifically focus on green and blue channels. To overcome limitations of DCP, Carlevaris-Bianco et al. [CME10] propose haze removal based on scene depth, derived from the strong difference in attenuation between the three image color channels in water. Emberton et al. [ECC15] presents a method that specifically avoids the typical oversaturation of veiling light, i.e., very bright regions in underwater images. An example of so-called fusion-based methods is presented by Ancuti et al. [AAHB12], obtaining enhanced underwater images via multiscale fusion of color corrected and contrast enhanced versions of the original underwater image, yielding reduced noise and improved global contrast.

Regarding image sequences, Drews et al. [DNCE15] propose a method to compute depth based on optical flow from pairs of directly successive frames. For more robust underwater depth estimation, Li et al. [LTT*15] apply stereo matching in combination with haze intensity estimation. Some research has also already focused on performance in order to enable real-time processing, e.g., by reducing model complexity [LAP20; IXS20], typically sacrificing some image quality. So far, only little focus was put on temporal stability for underwater video enhancement. Ancuti et al. [AAHB12] propose time-domain bilateral filtering to directly incorporate time-sequence information for better temporal coherence. Qing et al. [QYX*16] suggest to improve temporal color consistency by spatial–temporal information fusion for transmission and background light estimation, based on correlation between adjacent frames.

To demonstrate our approach, we will extend two single-frame methods based on monocular depth estimation, to enhance monocular RGB underwater video footage. The first method is by Peng et al. [PZC15]. Compared to similar work that uses the image formation model (IFM) by Jaffe [Jaf90] to enhance the contrast of underwater images, Peng et al. uses the blur of objects for depth estimation, which is often stronger for more distant objects due to light scattering. This allows objects of different distances to be weighted accordingly. Their approach achieves comparatively better results than other methods that build on the IFM. The al-

gorithm specifies three steps for creating the depth map. First, the pixel blur is estimated using the difference between the original image and the same image filtered several times through a Gaussian filter. Then, the depth map is generated using a maximum filter, assuming that in a small, local area depth remains uniform. In the last step, the guided filter of He et al. [HST12] is used, to refine the depth map by closing gaps using morphological reconstruction.

To complete the color correction, the depth information is used to estimate the background light from the most distant pixels. The correction is then applied by employing the simplified IFM for underwater images:

$$I(x) = J(x)t(x) + \beta(1 - t(x))$$

with J as scene intensity, β as background light, and t as transmission map (how much scene intensity reaches the camera).

As a second example, we implement an improved variant of the aforementioned work, proposed by Peng and Cosman [PC17]. This not only uses depth of field to determine the distance of underwater objects, but also incorporates light absorption, which is greater for more distant objects. Furthermore, the calculation of the background illumination is improved by not calculating it from the entire image, but from candidate points that originate from blurred regions of the image, i.e. the distant background.

In our work, we extend these methods by using optical flow data from successive frames to specifically improve temporal stability of the depth maps used to guide the image enhancement.

3. Method

In order to improve the perceived visual quality of underwater videos, depth information is employed by many algorithms, e.g., Peng et al. [PZC15] and Peng and Cosman [PC17]. The color value of individual pixels can then be corrected depending on the corresponding estimated depth. However, when used on video, effects such as motion blur can cause wildly different depth maps being computed for successive frames.

Our temporal stabilization approach is based on the assumption that a scene changes only slightly over short periods of time, and thus successive depth maps should provide similar values. However, when naively averaging different frames, the resulting depth map is heavily blurred due to video motion. Therefore, our approach is based on optical flow by Farnebäck [Far03] to combine multiple successive depth maps. Specifically, we compute per-frame optical flow, and track the x and y pixel positions of frame i (x_i, y_i) over several frames in the RGB video to use the shifted pixel positions in frame j

$$(x_j, y_j) = d_{j \leftarrow i}(x_i, y_i)$$

to average the depth maps, shifted accordingly via backward projection to avoid holes.

However, two limitations follow from this: First, for performance reasons we do not calculate the disparity (optical flow) from every frame to every other frame $d_{j \leftarrow i}$ in the chosen time window. Instead, we approximate it by accumulating the displacements over the span of images $(i, i+1, ..., j-2, j-1, j)$ as follows:

$$\begin{aligned}(x_j, y_j) &= d_{j \leftarrow j-1}(x_{j-1}, y_{j-1}) \\ &\textit{with} \\ (x_{j-1}, y_{j-1}) &= d_{j-1 \leftarrow j-2}(x_{j-2}, y_{j-2}) \\ &... \\ (x_{i+1}, y_{i+1}) &= d_{i+1 \leftarrow i}(x_i, y_i)\end{aligned}$$

Due to accumulation errors this is only an approximation. An example of this method is shown in Figure 2.

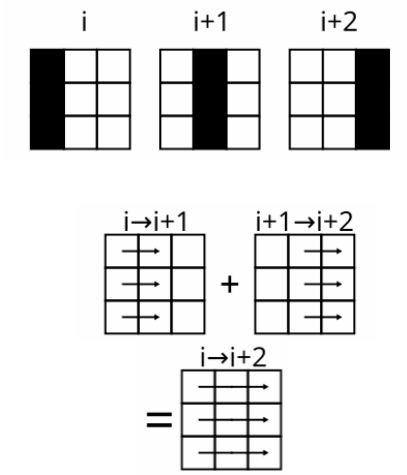

Figure 2: *Accumulative displacement $d_{i+2 \leftarrow i}$, combining two single step optical flow fields into one (bottom), for an exemplary 9×9 b/w image sequence (top).*

Second, as soon as a pixel moves outside the image area, its information can no longer be used in subsequent frames. To make the method more robust despite these limitations, we run the procedure in both time directions. Thus, for a chosen window of size n, frames are at most $n/2$ away from the current time step. Furthermore, a pixel that moves out of view in one time direction is likely to be in frame in the other direction.

Figure 3 shows the depth map of an exemplary frame (A), and cutouts showing the same image patch for multiple subsequent frames. As can be seen in form of a change in the average brightness, depth maps computed from single frames are consistent in themselves, but not stable over time (B). This instability can be strongly reduced, as demonstrated in the row below (C), showing the same cutout of depth maps computed for the same frames, but stabilized with our approach.

4. Evaluation

To test and compare our approach, we use the methods *Peng et al. 2015* [PZC15] and *Peng and Cosman 2017* [PC17], each in their original single frame version and with our extension to videos.

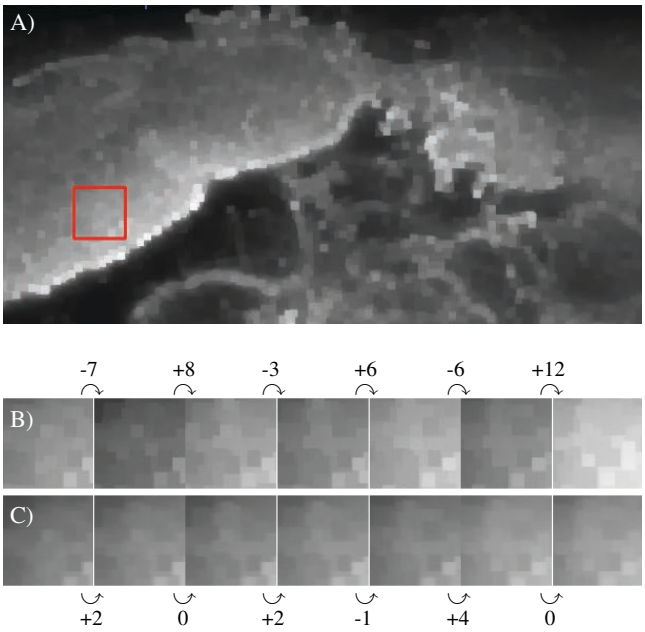

Figure 3: *Depth estimation for scene 1 (A). Cutouts of exemplary consecutive frames showing unstable depth without (B) and with our stabilization (C). Annotated numbers with curved arrows indicate mean brightness(depth) difference between successive frames. Visual contrast is slightly increased for illustration purposes.*

For the evaluation, we use eight video sequences of different lengths (see Table 1) showing different underwater scenes, as depicted in Figure 4. All videos were provided under a creative commons license with a resolution of at least 1920×1080 px.

The extension of the multi-frame algorithms is implemented using a window of nine frames, the four previous frames, the current frame, and the four succeeding frames. Additionally, for Peng and Cosman 2017 [PC17], the parameter $\mathrm{D}\infty$, which controls the overall strength of color correction, is adjusted to the scene. In the case of scene 5, this reduces the spontaneous occurrence of green saturated frames, in the other cases an unnatural red saturation.

Video	# Frames / Duration (s)	$\mathrm{D}\infty$
1	858 / 29	8 (default)
2	259 / 9	4
3	238 / 8	8 (default)
4	146 / 5	6
5	509 / 18	3
6	1359 / 48	8 (default)
7	426 / 15	8 (default)
8	1103 / 38	4

Table 1: *Test video lengths, along with the value for parameter ($\mathrm{D}\infty$) used in Peng and Cosman 2017.*

4.1. Scenes

In the following we describe our test scenes, the visual results after applying the color correction algorithms, and scene-specific challenges regarding temporal consistency in the depth and background light estimation. Representative images are shown in Figure 4.

Scene 1 shows a camera pan across the ocean floor with the sea on the horizon to a coral in which a ray is hiding. At first, therefore, the background is mainly blue due to the water and some creatures are visible in the lower half of the video. From about halfway through the video, water is barely perceptible, as filming now takes place in close proximity to the frame-filling coral.

After correction, the colors of the elements that can be seen closer in the video at the beginning stand out more strongly. As the video progresses, this color difference from the original are less noticeable as all the elements are at a similar distance from the camera and the blue turbidity of the water is only slightly noticeable. The challenge here is that there is a steep change in depth between fore- and background in the video.

Scene 2 shows a ray moving close to the seafloor between rock formations that is followed by the camera. The substrate is very bright and therefore reflects a lot of light coming in from above. More rocks can be seen in the background. In addition, the water is typically clouded blue.

After correction, both the ray and the rocks in the foreground stand out more compared to the original scene. However, the turbidity in the background does not change much. The challenge is a panning motion with changing extreme light situations (bright sand, dark rock).

Scene 3 also features a ray being followed by the camera. However, it looks like the scene was shot in an aquarium. The water is very clear even at a greater distance and has almost no blue turbidity. Therefore, the ray in the foreground is already very clear, i.e. without the influence of the water.

After correction, the scene shows hardly any changes to the original. However, the colors of the ray get strongly saturated, since it is alone in the image for most of the scene. The challenge is the correct estimation of background light with very low water turbidity and dark background.

Scene 4 shows a ray that is being tracked by the camera, moving at a slightly further distance. It is clearly visible, but the turbidity of the water affects the view of it somewhat.

After correction, since only the ray is visible in the scene, apart from other objects in the background, the color saturation of it is increased so that it stands out in the video compared to the original scene and the blue turbidity is slightly reduced in the direct view of it. The challenge here is filming upwards, having the water surface in frame, together with incoming sunlight overexposing that part of the image.

Scene 5 shows a close-up of a ray that is close to the bottom. The ray is spotted with color, unlike the rays from the previous scenes. The captured image has a noticeable blue haze, although water is

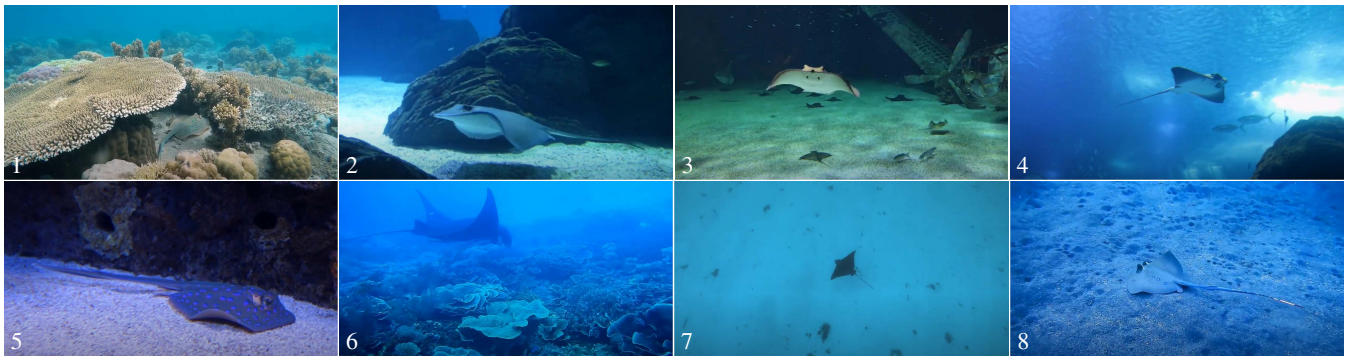

Figure 4: *Representative frames of the tested scenes 1–8 [wai16; Ehl19b; sus18; Ehl19c; Ehl19a; Fis18b; Joh18; Fis18a].*

not directly visible. The background is sandy and bright, and the background is very dark.

After correction, the whole scene is slightly oversaturated. The challenge is that there is no real depth in the scene, but all elements are very close to the camera, so that a background cannot be reliably determined with the method.

Scene 6 shows a shot of a more distant ray in very murky seawater. As a result, this can only be clearly seen by its silhouette. In contrast, the texture of its surface is barely perceptible. In the lower part, you can also see the seabed, on which there are various plants Of these, also, due to the strong turbidity of the water, hardly any colors can be perceived.

After correction, some elements from the seabed, whose distance from the camera is closer, are better perceivable. The visibility of the ray (in the background) will remain similarly difficult because it is at too great a distance from the camera. The challenge is restoring colors from extremely discolored input data.

Scene 7 shows a camera perspective perpendicular to the seabed on which a ray is moving. The bottom itself is very neutral, monochromatic and there are few other elements to be seen except for the ray. Due to the fact that the video was captured at a greater distance from the bottom, only the silhouette of the ray can be seen here as well, and not the actual texture of its body.

After correction, because of the light blue water turbidity and the distance from the elements in focus in the image, such as the ray, the image colors are changed throughout the entire frames and take on increased color saturation. The challenge is the flat floor with uniform distance to the camera, and the low amount of stationary features compared to the high portion that is moving.

Scene 8 again shows a shot of a ray, which is moving along the seafloor. Unlike the previous scenes, however, its body texture is clearly visible. The scene has a noticeable blue color cast, which occurs due to the absorption property of the water.

After correction, the texture of the ray as well as the seabed stand out in its color saturation. However, due to the fact that there are again few different distances in the scene, no clear separation between background and foreground can be made. The challenge here is the low visual contrast of the moving ray in front of the seafloor.

4.2. Evaluation Framework

For the evaluation of single frames we use the "Underwater Image Understanding" test suite by T. Krahn [KM16], which is an implementation based on the Underwater Image Quality Measures (UIQM) algorithm published by Panetta et al. [PGA15].

With this test suite we assess our method on the test scenes using two established metrics for underwater images: Underwater Image Colorfulness Measure (UICM), and Underwater Image Sharpness Measure (UISM). They evaluate different attributes of images, inspired by properties of the human visual system.

UICM evaluates the chromaticity of an image and is based on E. Hering's opponent theory [Her20]. This states that the human visual system does not perceive colors independently, but in pairs of opposite colors, so a value for evaluating colorfulness is calculated based on the following pairs: yellow–blue and red–green. The larger the calculated UICM value for an image, the greater the image's colorfulness.

UISM evaluates the sharpness of an image. Underwater, forward scattering effects occur which lead to the degradation of image sharpness in underwater images. A UISM value is calculated by applying a Sobel edge detector to the image and then evaluating the sharpness of the edges for the resulting edge map. The greater the sharpness of an image, the greater the UISM value.

Because the metrics evaluate individual frames, it is not possible to evaluate the resulting videos as a whole. However, it does allow to compare frames of the videos generated with Peng et al. 2015 and Peng and Cosman 2017, with and without our extension.

All 8 video sequences are split up into individual frames, which are processed by the four approaches (two methods, with/without extension). Finally, the scores for both metrics are computed for the original and all processed frames, yielding for the following 5 conditions:

- ***original***: Original frames
- ***single15***: Optimized with Peng et al. 2015 [PZC15]
- ***multi15* (ours)**: Opt. with Peng et al. 2015 + extension
- ***single17***: Optimized with Peng and Cosman 2017 [PC17]
- ***multi17* (ours)**: Opt. with Peng and Cosman 2017 + extension

4.3. Results

The results of our evaluation are summarized in Table 2, showing one plot per scene (1–8) and metric (UICM, UISM). For each of the plots, the per-frame (x-axis) score of the respective metric (y-axis) is shown for the five methods listed above. It should be noted that the goal was not to improve the UICM or UISM values but to smoothen them out, as sudden changes in the metrics result in visible flickering in the videos. We would also like to mention that for the set of scenes tested, the older *single15* method does not necessarily produce worse results compared to the newer *single17* method.

Coloring (UICM) In each evaluated scene, both the single frame *single15* method and the multi-frame *multi15* method have a comparably higher chromaticity level than the corresponding original frame. It can be seen that our multi-frame method hardly differs from the single-frame method when considering this metric, and is only slightly weaker compared to it. Compared to this, the *single17* method shows weaknesses in some scenes, as the results are close to those of the original frame in scenes 2, 6, and 7 almost across the entire scene. In addition, it can be seen that the chromaticity of the *multi17* method performs always equal to or better than *single17*, except for areas of overly strong fluctuation, e.g., in scene 7. Between the respective single-frame methods, we can also see that the chromaticity in *single15* is typically above the result of *single17*, except for scene 8 and parts of scene 1 and 5. The same relationship can also be seen between our multi-frame methods. Here, there is only one opposite case in scene 8, in which the *multi17* method has a higher average chromaticity value than the *multi15* method.

As can be seen in the UICM metric in Table 2 and in the enlarged excerpt from scene 1 in Figure 5, the main benefit of our multi-frame extension is the smoothing of spikes in the graph. These spikes correspond to perceived flickering in the videos caused by quickly changing color corrections calculated by the single-frame methods.

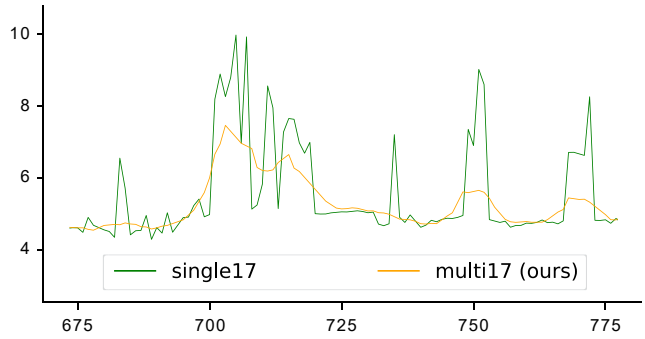

Figure 5: *UICM for scene 1: Enlarged example excerpt of single17 and multi17 (ours) to showcase the smoothing effect of our method.*

Image sharpness (UISM) In scenes 1-3, 6, and 8, the image sharpness of all four methods is comparable to that of the original video. In scene 1, all four methods initially offer slightly improved image sharpness compared to the original for about half the video's duration. However, after that point, *single17* and *multi17* start to exhibit significant fluctuations, oscillating around the sharpness value of the original video. A similar behavior can be observed in scene 7. In scene 3, the *single15* and *multi15* method are initially slightly less sharp than the original, but from 1/3 of the scene onward, all methods can improve sharpness of the frames. For scenes 5, 7, and 8, all methods yield consistently improved image sharpness above the original frame, throughout the full scenes.

For scenes 4 and 7, a different trend can be seen. On average, all four methods are clearly above the image sharpness value of the original frames, but there are dips in regular, temporal intervals of about 12 frames, in which the sharpness of all four methods falls to around the image sharpness value of the original video.

From all the scenes, it can be deduced that the image sharpness is not negatively affected by the four methods, but rather improved. It can be seen that the multi-frame extension does not negatively affect image sharpness (compared to the single-frame methods), e.g., does not introduce disturbing image artifacts concerning image sharpness.

4.4. Limitations

Our method is limited by the precision of the optical flow. As the resulting flow fields usually contain imperfections, so does the color correction. This is especially true for videos where objects move quickly between different positions, where optical flow analysis often fails.

In addition, our method is still dependent on the quality of the videos being processed. This means that unstable, poorly exposed, or blurry videos will in turn lead to poor results. The reason for this is that important information about the depth of elements potentially cannot be calculated or can only be calculated incorrectly in the case of blurred frames, leading to problems in extracting the correct background illumination color. Conversely, well-exposed, depth-rich and sharp video of underwater scenes enables greater and more accurate improvements compared to the original considering the colorfulness and sharpness of the image. Methods implementing our extension can benefit the most for underwater videos suffering from sudden disturbances, e.g., brightness jumps due to faulty automatic exposure.

Regardless of the input quality, it is not possible in the current CPU-only implementation to process the video in real-time. The individual steps, the single frame method, especially the optical flow calculation, but also the recombination of the successive frames, require too much processing time. In average over the 8 tested videos, the full processing pipeline takes about 1.55 s per frame ($\sigma = 0.027$ s) for the *multi15* method and 0.99 s per frame ($\sigma = 0.016$ s) for the *multi17* method. Of the total processing time, the proposed extension accounts for about 83 % (*multi15*) to 76 % (*multi17*).

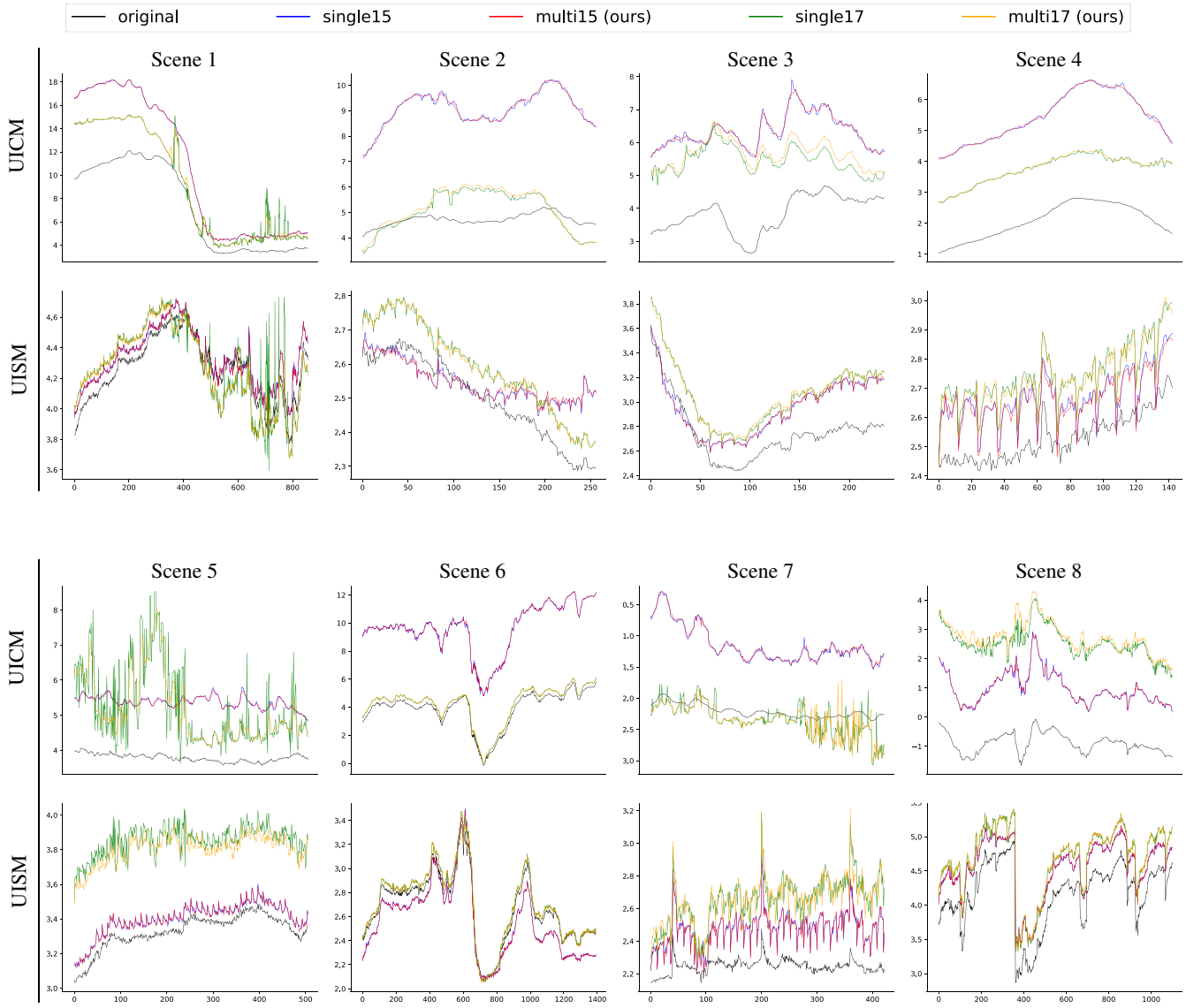

Table 2: *UICM and UISM metrics scores (y-axis) plotted per frame (x-axis) for each condition (*original*,* single15*, Peng et al. 2015 + extension,* single17*, Peng and Cosman 2017 + extension) and all 8 tested scenes.*

5. Conclusion

In conclusion, we have presented an approach to extend single-frame depth estimation-based underwater color enhancement methods to video, for improved temporal consistency. The extension was applied to two example implementations of such single-frame approaches and was evaluated using dedicated underwater image quality metrics. The results of our evaluation show that our proposed multi-frame extension successfully improves temporal consistency and even strengthens colorfulness and image sharpness.

Possible future improvements include real-time processing, which is currently primarily limited by the optical flow computation. This would allow for interactive usage for post-processing underwater scenes in video editing. Alternative to our approach of using multi-frame information to stabilize depth that was computed from single (pairs of successive) frames, an interesting approach might be to directly infer more stable depth using information from multiple frames e.g. through structure-from-motion techniques. Further developments of general extensions to underwater image enhancement methods include the detection and reduction of unnatural color saturation, and slower flickering that is not detectable with a frame-to-frame approach. Lastly, our approach to stabilize depth maps using optical flow potentially could be used for other depth guided image enhancement methods.

6. Acknowledgments

Partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 491805996 and by the Lower Saxony Ministry of Science and Culture under grant number ZN3994 within the Lower Saxony "Vorab" of the Volkswagen Foundation and supported by the Center for Digital Innovations (ZDIN).

References

[AAHB12] ANCUTI, COSMIN, ANCUTI, CODRUTA ORNIANA, HABER, TOM, and BEKAERT, PHILIPPE. "Enhancing underwater images and videos by fusion". *2012 IEEE conference on computer vision and pattern recognition*. IEEE. 2012, 81–88 2.

[CC12] CHIANG, JOHN Y and CHEN, YING-CHING. "Underwater image enhancement by wavelength compensation and image dehazing". *Proc. IEEE J. on IP* 21.4 (2012) 2.

[CME10] CARLEVARIS-BIANCO, NICHOLAS, MOHAN, ANUSH, and EUSTICE, RYAN M. "Initial results in underwater single image dehazing". *Oceans 2010 Mts/IEEE Seattle*. IEEE. 2010, 1–8 2.

[DNCE15] DREWS, PAULO, NASCIMENTO, ERICKSON R, CAMPOS, MARIO FM, and ELFES, ALBERTO. "Automatic restoration of underwater monocular sequences of images". *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2015, 1058–1064 2.

[DNM*13] DREWS, PAUL, NASCIMENTO, ERICKSON, MORAES, FILIPE, et al. "Transmission estimation in underwater single images". *Proceedings of the IEEE international conference on computer vision workshops*. 2013, 825–830 2.

[ECC15] EMBERTON, SIMON, CHITTKA, LARS, and CAVALLARO, ANDREA. "Hierarchical rank-based veiling light estimation for underwater dehazing". *Proceedings of the British Machine Vision Conference (BMVC)*. 2015, 125.1–125.12 2.

[Ehl19a] EHLERS, MAGDA. *A Blue Spotted Sting Ray In An Aquarium*. Pexels. 2019. URL: https://www.pexels.com/video/a-blue-spotted-sting-ray-in-an-aquarium-2561847 5.

[Ehl19b] EHLERS, MAGDA. *A Sting Ray On The Move*. Pexels. 2019. URL: https://www.pexels.com/video/a-sting-ray-on-the-move-2556843 5.

[Ehl19c] EHLERS, MAGDA. *Underwater Footage*. Pexels. 2019. URL: https://www.pexels.com/video/underwater-footage-2556894 5.

[Far03] FARNEBÄCK, GUNNAR. "Two-frame motion estimation based on polynomial expansion". *Image Analysis: 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, June 29–July 2, 2003 Proceedings 13*. Springer. 2003, 363–370 3.

[FC20] FU, XUEYANG and CAO, XIANGYONG. "Underwater image enhancement with global–local networks and compressed-histogram equalization". *Signal Processing: Image Communication* 86 (2020), 115892 2.

[FIS18] FABBRI, CAMERON, ISLAM, MD JAHIDUL, and SATTAR, JUNAED. "Enhancing underwater imagery using generative adversarial networks". *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, 7159–7165 2.

[Fis18a] FISK, TOM. *A Stingray Under Blue Waters*. Pexels. 2018. URL: https://www.pexels.com/video/a-stingray-under-blue-waters-1312355/ 5.

[Fis18b] FISK, TOM. *Manta Ray Swimming Above Coral Reef*. Pexels. 2018. URL: https://www.pexels.com/video/manta-ray-swimming-above-coral-reef-1151322 5.

[Her20] HERING, EWALD. *Grundzüge der lehre vom lichtsinn*. Springer, 1920 5.

[HST12] HE, KAIMING, SUN, JIAN, and TANG, XIAOOU. "Guided image filtering". *IEEE transactions on pattern analysis and machine intelligence* 35.6 (2012), 1397–1409 3.

[HWZ*22] HU, KAI, WENG, CHENGHANG, ZHANG, YANWEN, et al. "An overview of underwater vision enhancement: from traditional methods to recent deep learning". *JMSE* 10.2 (2022), 241 2.

[IXS20] ISLAM, MD JAHIDUL, XIA, YOUYA, and SATTAR, JUNAED. "Fast underwater image enhancement for improved visual perception". *IEEE Robotics and Automation Letters* 5.2 (2020), 3227–3234 2.

[Jaf90] JAFFE, JULES S. "Computer modeling and the design of optimal underwater imaging systems". *IEEE Journal of Oceanic Engineering* 15.2 (1990), 101–111 2.

[Joh18] JOHNSTON, SEAN. *A Sting Ray Moving On The Sea Floor*. Pexels. 2018. URL: https://www.pexels.com/video/a-sting-ray-moving-on-the-sea-floor-3106428 5.

[KM16] KRAHN, T. and MÖLLER, T. *Underwater Image Quality Measure (UIQM)*. https://github.com/tkrahn108/UIQM. Accessed: 2023-05-30. 2016 2, 5.

[LAP20] LI, CHONGYI, ANWAR, SAEED, and PORIKLI, FATIH. "Underwater scene prior inspired deep underwater image and video enhancement". *Pattern Recognition* 98 (2020), 107038 2.

[LLU*18] LU, HUIMIN, LI, YUJIE, UEMURA, TOMOKI, et al. "Low illumination underwater light field images reconstruction using deep convolutional neural networks". *Future Generation Computer Systems* 82 (2018), 142–148 2.

[LSEJ17] LI, J., SKINNER, KA., EUSTICE, RM., and JOHNSON-ROBERSON, M. "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images". *IEEE Robotics and Automation letters* 3.1 (2017), 387–394 2.

[LTT*15] LI, ZHUWEN, TAN, PING, TAN, ROBBY T, et al. "Simultaneous video defogging and stereo reconstruction". *IEEE conference on computer vision and pattern recognition*. 2015, 4988–4997 2.

[PC17] PENG, YAN-TSUNG and COSMAN, PAMELA C. "Underwater image restoration based on image blurriness and light absorption". *IEEE transactions on image processing* 26.4 (2017), 1579–1594 2–5.

[PGA15] PANETTA, KAREN, GAO, CHEN, and AGAIAN, SOS. "Human-visual-system-inspired underwater image quality measures". *IEEE Journal of Oceanic Engineering* 41.3 (2015), 541–551 2, 5.

[PZC15] PENG, YAN-TSUNG, ZHAO, XIANGYUN, and COSMAN, PAMELA C. "Single underwater image enhancement using depth estimation based on blurriness". *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2015, 4952–4956 2, 3, 5.

[QYX*16] QING, CHUNMEI, YU, FENG, XU, XIANGMIN, et al. "Underwater video dehazing based on spatial–temporal information fusion". *Multidimensional Systems and Signal Processing* 27 (2016), 909–924 2.

[sus18] SUSIESMITHUK. *Aquarium Manta Strahl*. Pixabay. 2018. URL: https://pixabay.com/de/videos/aquarium-manta-strahl-unterwasser-13833 5.

[wai16] WAIGUOBOX. *Stachelrochen Fische Marine*. Pixabay. 2016. URL: https://pixabay.com/de/videos/stachelrochen-fische-marine-strahl-5271 5.

[WZCW17] WANG, Y., ZHANG, J., CAO, Y., and WANG, Z. "A deep CNN method for underwater image enhancement". *International conference on image processing*. IEEE. 2017, 1382–1386 2.

[ZPIE17] ZHU, J, PARK, TAESUNG, ISOLA, PHILLIP, and EFROS, ALEXEI A. "Unpaired image-to-image translation using cycle-consistent adversarial networks. CoRR abs/1703.10593 (2017)". *arXiv preprint arXiv:1703.10593* 3 (2017) 2.

[ZYZ23] ZHOU, JINGCHUN, YANG, TONGYU, and ZHANG, WEISHI. "Underwater vision enhancement technologies: A comprehensive review, challenges, and recent trends". *Applied Intelligence* 53.3 (2023), 3594–3621 2.