A motion compensated filtering approach to remove sunlight flicker in shallow water images

Nuno Gracias*, Shahriar Negahdaripour†, Laszlo Neumann*, Ricard Prados*, Rafael Garcia*

* Underwater Vision Lab	† Underwater Vision and Imaging Lab
Computer Architecture and Technology Department	Electrical and Computer Engineering Department
University of Girona	University of Miami
Avda. Lluis Santalo S/N, 17071 (Girona)	Coral Gables, FL 33124–0640
Email: {ngracias, rprados, lneumann, rgarcia}@eia.udg.edu	Email: shahriar@miami.edu

Abstract— **A common problem in video surveys in very shallow waters is the presence of strong light fluctuations, due to sun light refraction. Refracted sunlight casts fast moving patterns, which can significantly degrade the quality of the acquired data.**

Motivated by the growing need to improve the quality of shallow water imagery, we propose a method to remove sunlight patterns in video sequences. The method exploits the fact that video sequences allow several observations of the same area of the sea floor, over time. It is based on computing the image difference between a given reference frame and the temporal median of a registered set of neighboring images. A key observation is that this difference will have two components with separable spectral content. One is related to the illumination field (lower spatial frequencies) and the other to the registration error (higher frequencies). The illumination field, recovered by lowpass filtering, is used to correct the reference image. In addition to removing the sunflickering patterns, an important advantage of the approach is the ability to preserve the sharpness in corrected image, even in the presence of registration inaccuracies.

The effectiveness of the method is illustrated in image sets acquired under strong camera motion containing non-rigid benthic structures. The results testify the good performance and generality of the approach.

I. INTRODUCTION

Taking an appealing photograph of the seafloor in shallow waters seems an easy task. However, on a sunny day, this photograph will include not only the reflectance of the seafloor, but also sun flickering artifacts. These artifacts are caused by the intersection of the sunlight rays with the water surface waves, and appear in the image as bright stripes that change quickly in both space and time. Figure 1 illustrates this phenomenon. Sun flicker corrupts image appearance, altering human perception of the scene and affecting the behavior of image processing algorithms. Thus, development of techniques to reduce or eliminate these artifacts becomes crucial in order to ensure optimal performance of underwater imaging algorithms.

II. BACKGROUND AND RELATED WORK

The physical and statistical properties of light fields underwater have been extensively studied in the past [1], [2], [3]. However, somehow unexpectedly, the problem of reducing the effects of refracted sunlight has not been suitably addressed in the context of underwater video processing. Previous efforts have concentrated on using homomorphic filtering [4] and
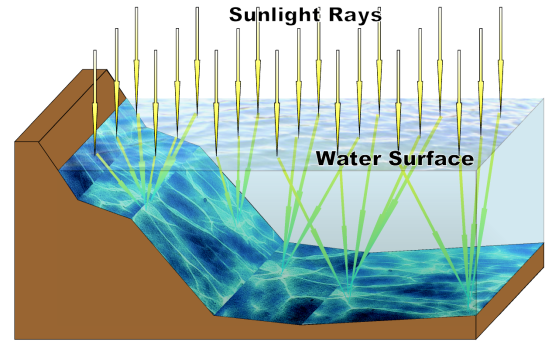

Fig. 1. Refracted sunlight creates illumination patterns on the seafloor, which vary in space and time following the dynamics of surface waves.

histogram equalization for the case of a single image, which is essentially an ill-posed problem.

Garcia *et al* [5] compare existing methodologies to deal with low-contrast, nonuniform illumination in underwater image sequences. The reviewed techniques include local histogram equalization, homomorphic filtering, and subtraction of the illumination field. The homomorphic filtering assumes a multiplicative composition of illumination and reflectance fields, while the subtraction of the illumination assumes additive. Results include a shallow water example with frame–to–frame illumination variations due to sunlight. The processing is done independently for each frame, therefore not using temporal correlations. Adaptive local histogram equalization (CLAHE [6]) or specification (CLAHS [7]) improves the appearance of the image by promoting similar levels of contrast over all areas of the image. However it tends to amplify image noise in low–contrasted areas. Homomorphic filtering (when applied to a single image) implicitly assumes that the illumination and reflectance are separable in the frequency domain, where illumination will be a low–frequency field. When that assumption is not valid, then the low frequency content of the reflectance field is affected by the filtering. Motivated by the reduction of seams in mosaics, Rzhanov and Gu [8] use the median image of a set of frames to reduce brightening and darkening trends of individual images. A second order polynomial fit is applied to model very smooth irradiance changes of individual frames.

This fit is replaced with a similar fit obtained from the median image. As formulated, this approach it is not intended to deal with the complexity of sun flickering patterns.

Schechner and Karpel [9] present a method for attenuating sunlight flicker patterns, inspired by effective methods for shadow removal in open–air scenes [10], [11]. This approach is based on the observation that the spatial intensity gradients of the caustics tend to be sparse (*i.e.* affect small areas of the image). Therefore, by performing the temporal median over the gradients of a small number of images, it is possible to obtain a gradient field where the effects of the varying illumination are eliminated or at least greatly reduced. The flicker–free image is obtained by integrating the median gradient field. As the gradient field is likely to be inconsistent, a least squares approximation is performed using reversed derivative kernels. This approach does not attempt to estimate any particular time instance of the refracted sunlight illumination field. Rather, the flicker–free image is reconstructed using contributions from all images, as determined by the temporal median. Therefore it is not clear how this method behaves under camera motion, where small registration errors are practically unavoidable. This method was implemented and compared in the results section of this paper.

III. APPROACH

Let us consider a set of registered images. We refer to a given image by the discrete parameter i which indexes the images temporally. The radiance value L of a given pixel with coordinates (x, y) can be modelled as

$$L_i(x, y) = E_i(x, y) \cdot R_i(x, y)$$

where E_i is the irradiance of the sunlight over the 3D scene at the location defined by pixel (x, y) at time i, after absorption in the water. $R(x, y)$ is the bidirectional reflectance distribution function. For simplicity, we assume a scene where diffuse reflectance models are applicable. R is a constant, independent of both light direction and view direction. If the registration is perfect, $R_i(x, y)$ will constant in time since (x, y) will correspond to the same elementary region of the 3D scene for all image.

Let us convert the above equation to a logarithmic scale. This allows the use linear filtering over the illumination and reflectance. After such operations, we apply the exponential function to recover the linear scale. Using lower case to denote log values (e.g. $l = \log L$),

$$l_i(x, y) = e_i(x, y) + r_i(x, y) \quad (1)$$

Let us consider the stack of temporally changing terms of Eq. (1) over a pixel (x, y). For an infinite sequence the average of the irradiance would be constant over the entire surface, in the case of homogenous water and constant depth. For realistic finite cases, the median significantly converges faster to this constant than the average value [12]. Let us considering the median of the radiances over an interval $i \in [i_0, i_1]$

$$l_{med}(x, y) = med_{[i_0, i_1]} l_i(x, y)$$

For a sufficiently large interval, we can assume that it contains a practically constant irradiance (i.e. spatially independent) and a reflectance term dependent on (x, y),

$$l_{med}(x, y) = med_{[i_0, i_1]} l_i(x, y) \approx e + r_{med}(x, y)$$

where $r_{med}(x, y)$ stands for an approximation to the median of reflectance (unobserved). This reflectance term is affected by the registrations errors. In the presence of small registration errors it tends to be a blurred version of the real reflectance, differing mainly in the absence of high spacial frequencies.

We now consider the difference $d_i(x, y)$ of a given image $l_i(x, y)$ with the median radiance $l_{med}(x, y)$. To the extent that the above approximation is valid, we can write

$$\begin{aligned} d_{l_i}(x, y) &= l_i(x, y) - l_{med}(x, y) \\ &\approx (e_i(x, y) - e) + (r_i(x, y) - r_{med}(x, y)) \end{aligned}$$

This difference $d_{l_i}(x, y)$ will have two components. The first component

$$d_{e_i}(x, y) = e_i(x, y) - e$$

relates to the instant fluctuation of the illumination field with respect to the constant irradiance. This component will have positive values in the over–illuminated areas where there is convergence of the refracted sunlight and will have negative values in the areas where the sunlight is diverted away. The second component

$$d_{r_i}(x, y) = r_i(x, y) - r_{med}(x, y)$$

relates to the difference between the reflectance of image i and the approximation to the median reflectance. For limited errors on the registration $r_{med}(x, y)$ will describe the lower frequencies of the reflectance field, and $d_{r_i}(x, y)$ will contain mainly high frequencies of the registration residues.

The key observation in our approach is that $d_{e_i}(x, y)$ and $d_{r_i}(x, y)$ have very distinct spectral content. For practical cases, the higher spacial frequencies of the term $d_{e_i}(x, y)$ are significantly lower that the lower frequencies of $d_{r_i}(x, y)$. This observation makes it feasable to separate the components by adequate low–pass filtering. Considering

$$\widetilde{d}_{l_i}(x, y) = lowpass\{d_{l_i}(x, y)\}$$

we assume

$$\widetilde{d}_{l_i}(x, y) \approx d_{e_i}(x, y)$$

The $\widetilde{d}_{l_i}(x, y)$ can be considered as an illumination correction field. An illumination corrected image $\widehat{L}_i$ is obtained by

$$\widehat{L}_i(x, y) = L_i(x, y) / \exp\left(\widetilde{d}_{l_i}(x, y)\right).$$

IV. IMPLEMENTATION

The implementation of the proposed method is summarized in the following steps

1) For all images L_i compute the logarithm of pixel intensities $l_i = \log(L_i)$.
2) For a given image l_i, consider the set $S_i = \{l_{i-k}, ..., l_i, ..., l_{i+k}\}$ of $2k+1$ time neighboring images centered in i.

3) Using robust feature matching [13], [14], find the 2D image registration parameters that relate all images of S_i with l_i.
4) Compute $l_{i_{med}}$ as the temporal median of image stack after projecting all images into the reference frame of l_i.
5) Compute $d_i = l_i - l_{i_{med}}$
6) Apply a low–pass filter to obtain

$$\widetilde{d}_i = lowpass\,\{d_i\}$$

7) Compute the illumination corrected image as

$$\widehat{l}_i = l_i - a \cdot \widetilde{d}_i + b$$

where a and b are found by minimizing a linear least squares criteria

$$(a,b) = \arg\min_{a,b} \left(d_i - a \cdot \widetilde{d}_i + b \right)^2$$

8) Invert the logarithm to recover the linear pixel intensities $\widehat{L}_i = \log^{-1}\left(\widehat{l}_i \right)$.

For the results presented in this paper the above steps are applied independently over each color channel. We assume the knowledge of the gamma values for each color channel. For unknown gamma values one can apply blind gamma estimation. An efficient method is described in [15], [16], which exploits the fact that gamma correction introduces specific higher-order correlations in the frequency domain. Having the gamma values we transform the intensities to linear scale. After the de–flickering the final output images are transformed into the sRGB space with the prescribed gamma value.

An underlying objective of this work is to obtain a fast method that can be applied large sets of images in useful time. The choice of using the median instead of the much faster average, is justified by the resilience of the median to strong caustics. Strong caustics lead to overexposure and intensity clipping in one or more of the color channels resulting in chromaticity changes in the original images. These clippings typically affect different regions of the images over time, given the non–stationary nature of the caustics. The median is not affected by such transient clippings, whereas the average will. For the cases where overexposures are unlikely to occur (for example when using high dynamic range cameras, or in deeper water where the flicker amplitude is smaller), then the average can be used instead of the median, thus making the processing faster. Another possibility (however slower) is the use a method based on the median of gradients [9].

As mentioned before, the spatial frequency content of the image difference d_i has two typically very separable components. The low frequency content of the illumination field allows for an important speed–up in the processing. The median image $l_{i\ med}$ can be computed at a considerable lower resolution than the resolution of the original images. The results in this paper were obtained by computing the median at a quarter resolution both horizontally and vertically, with important savings in processing time.

The low pass filtering is performed using a forth order Butterworth filter [17], with a cutoff frequency of $\omega = 0.05$, which was found to be adequate for all the image sequences tested.

The illumination correction in step 7 is performed by rescaling the values of filtered difference before applying to the original image. The estimation of scalars a and b can be avoided by setting a to the inverse of the DC gain of the filter and b to zero. However, slightly better results can obtained by estimating a and b from the data, specially if the gamma correction is not precisely known. The estimation implies solving a overdetermined set of linear equations (one per element of d_i) which can accelerated using subsampled versions of d_i and $\widetilde{d}_i$.

The intermediate results of the method are illustrated in Figure 2. This example was constructed from stack of nine registered frames over a relatively flat area, however containing some 3–D structure (rocks and algae). A low resolution median was constructed, and then subtracted from the original image. The two components of the difference image are easily seen. They can be effectively separated into the low frequency illumination field and the high frequency residues (resulting from imprecise registration and the low resolution median).

V. Experiments and Results

For the results in this paper, we have selected a number of distinct sequences acquired at different depths. The sequences were obtained from video surveys in shallow waters [18], [19] where the cameras are moving at rates of about 0.5 m/s. The images were corrected for lens distortion. To remove the sunlight flickering, the images were registered pair-wise and sequentially in time. The inter–frame motion was modeled by a simple planar–projective mapping with 8 degrees of freedom. This assumes an approximately flat environment, thus leading to small but noticeable registration errors for the sequences with 3D content.

Figure 3 show the original and corrected frames for very shallow waters of 2–3 meter depth, where the camera was close to the surface. Results for deeper waters (about 5 meter depth) are presented in Figure 4. These examples show the ability of the method to remove caustics without affecting image detail.

Figure 5 illustrates the impact of the sunflickering removal on a section of a mosaic of a ship grounding survey.

A. Comparison with median of gradients

For comparison purposes we implemented a sunflicker removal approach based on the median of gradients [11], [9]. For this we followed closely the algorithm provided in [9] namely in computing image derivatives on the natural logarithm of image intensities, pixel-wise temporal median among the stack of contributing images, establishing boundary conditions on the gradient fields before intensity recovery, and same color handling on the finally obtained intensity channels. The recovery of image intensities from gradients was done using a Poisson solver with Neumann boundary
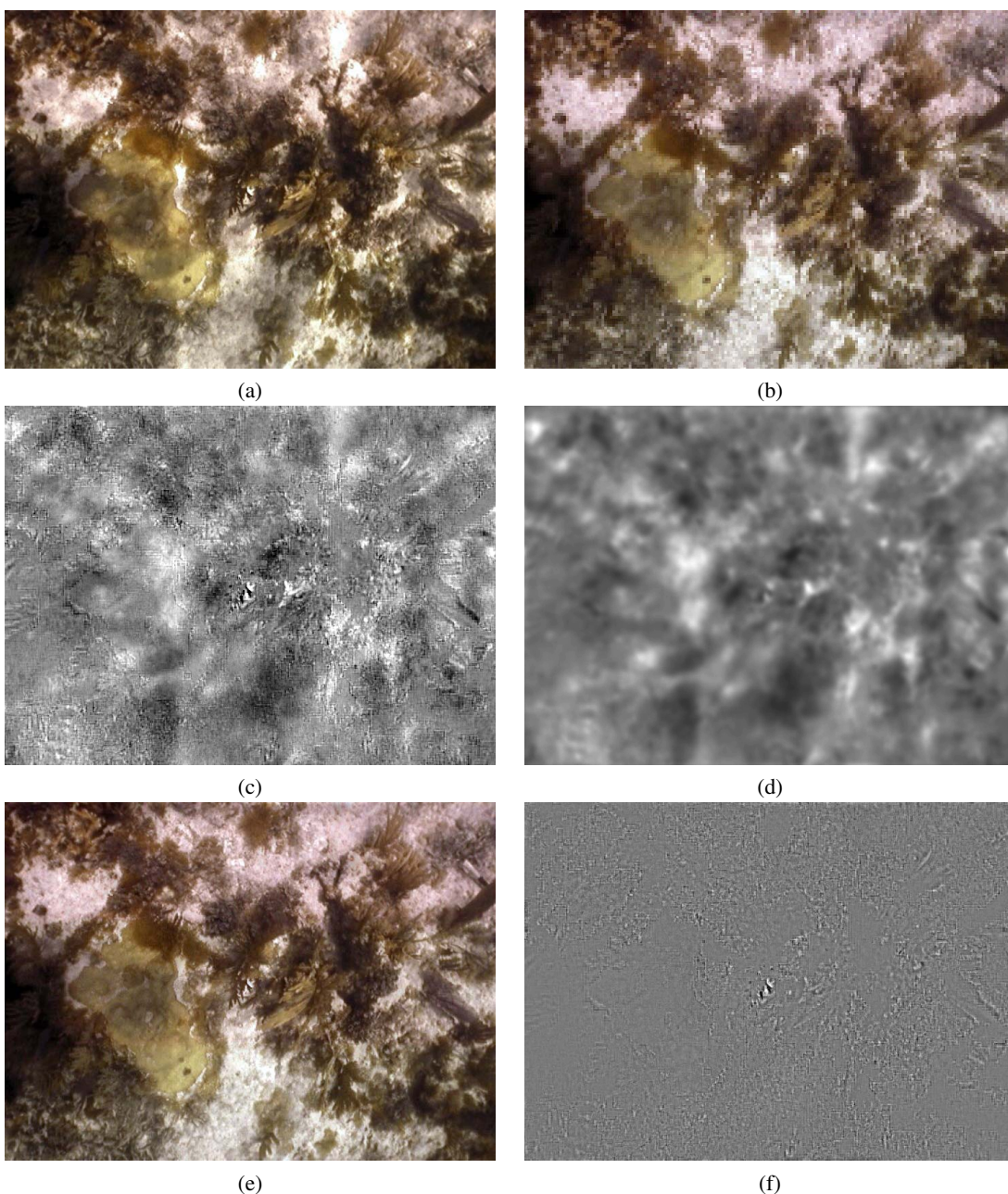
Fig. 2. Example of application of the motion compensated filtering algorithm — Original (reference) frame (a), sub–sampled temporal median of the registered stack after projection into the reference frame (b), image difference between the median image and the original for one of the color channels(c), low–pass filtered version of the difference (d), result of the illumination correction to the original image (e), and difference between the corrected image and the median (f). The values of images (c), (d) and (f) have been stretched to improve visualization.

conditions [20], mirroring that of [11]. The gradient median method is particularly effective in dealing with small sets of images with sharp transitions between lighted and shadowed regions. It uses contributions for the gradients of all images equally (without singling out a particular frame). As such, its use becomes limited in cases where the registration is not perfect. Figure 6 provides and example using 5 registered images where the 3D content of the scene is clearly not well represented by planar image motion. Expectedly, both the temporal median and the method based on the median of

gradients are affected, leading to blurry results.

VI. CONCLUSIONS

This paper presented a method for removing the illumination inconsistencies from video sequences under strong refracted sunlight. The key observation is that the illumination patterns have very low spatial frequency, with respect to the image resolution of standard underwater cameras used in surveying or mapping applications. The low frequency nature allows for very effective estimation of the required illumination correction, which can be adequately performed
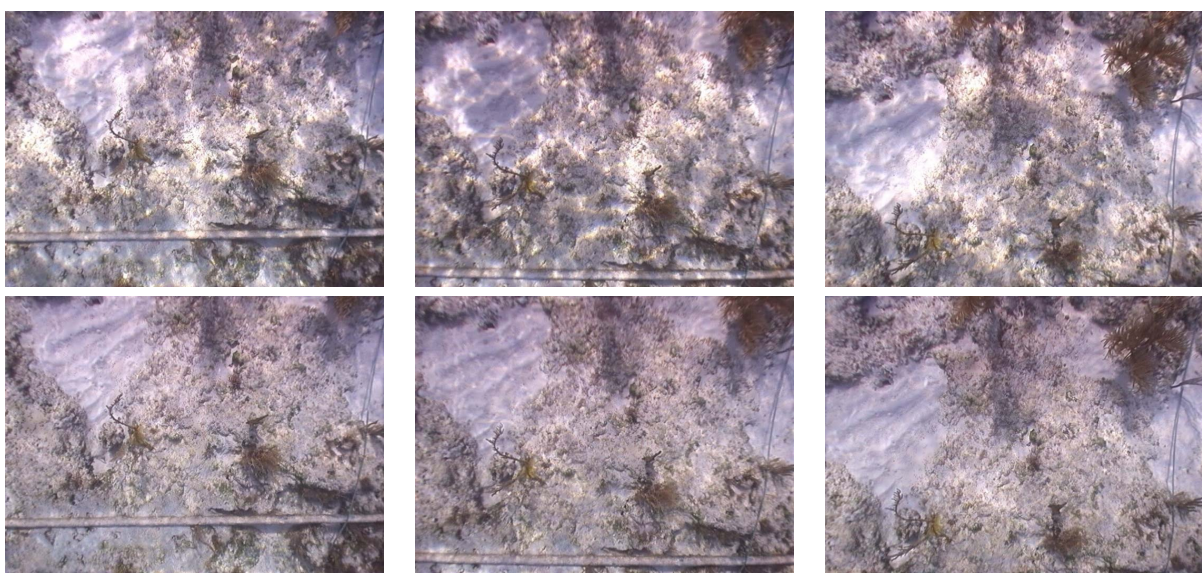
Fig. 3. Original frames of a shallow water sequence (2–3 meter) with refracted sunlight under camera motion (upper row) and resulting frames after applying our filtering technique (lower row).
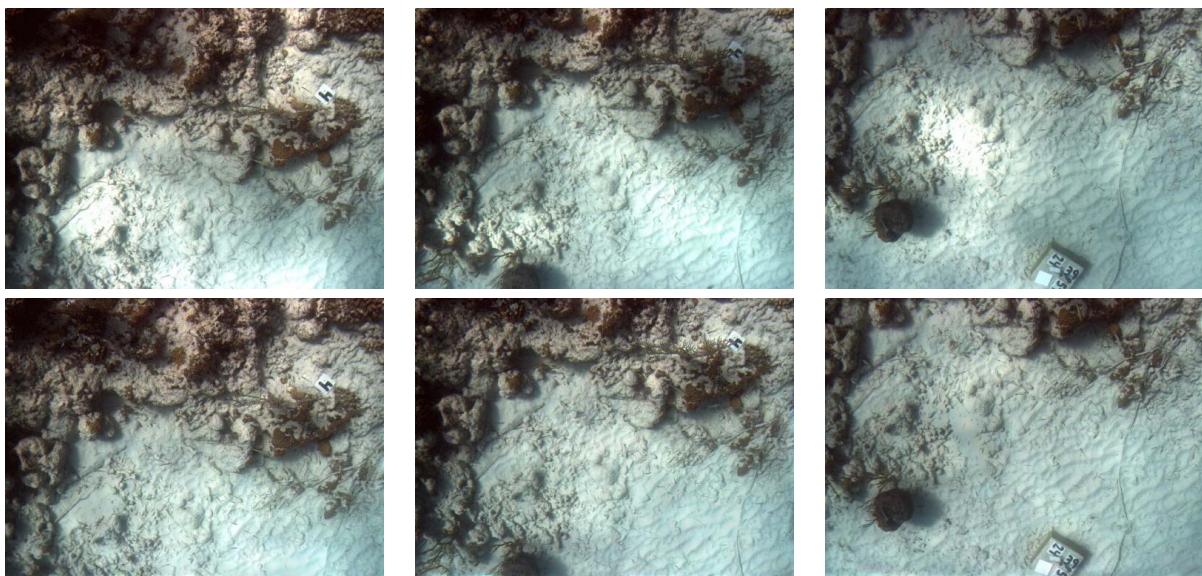

Fig. 4. Original frames of a deeper water sequence (approximately 5 meter) with intensity clippings due to strong caustics (upper row) and result of our filtering technique (lower row).

in the realistic cases of significant camera motion and limited registration accuracy.

Several results are presented for different camera and different water depth, under strong camera motion and non-rigid benthic content. The results testify the good performance and generality of the approach.

REFERENCES

[1] R. E. Walker, editor. *Marine light field statistics*. John Wiley, New York, 1994.
[2] A. Fraser, R. Walker, and F. Jurgens. Spatial and temporal correlation of underwater sunlight fluctuations in the sea. *IEEE Journal of Oceanic Engineering*, QE–5:195–198, 1980.
[3] V. Yakubenko and V. Nikolayev. Experimental studies of brightness fluctuations of the underwater light field. *Izvestiya, Atmosferic and Oceans Physics*, 14(8):624–628, 1978.
[4] A. Oppenheim, R. Schafer, and T. Stockham Jr. Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE*, 56(8):1264–1291, Aug. 1968.
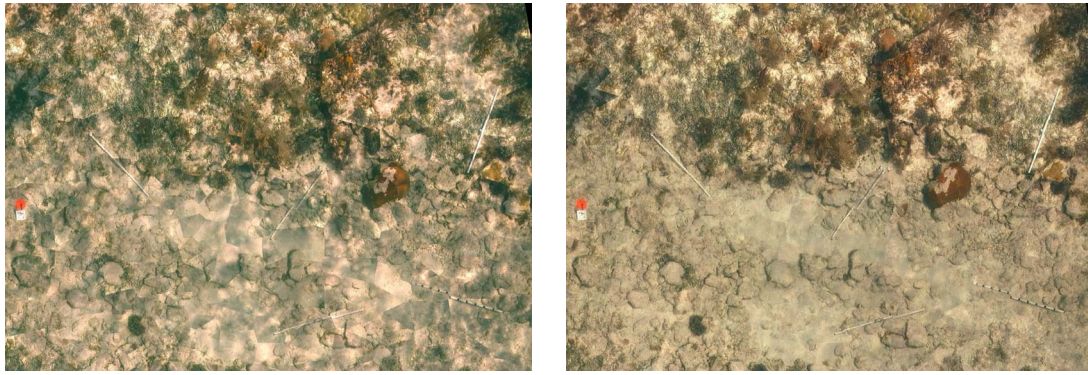
Fig. 5. Section of a mosaic from a ship grounding mosaic, illustrating the advantage of removing the flickering of the refracted sunlight. The left image was created with the original frames and presents a large number of visual artifacts due to illumination inconsistency. The right image was created with the filtered sequence and allows an easier interpretation of the benthic structures. Both images were created using a simple blending method of selecting the contribution from the closest image.
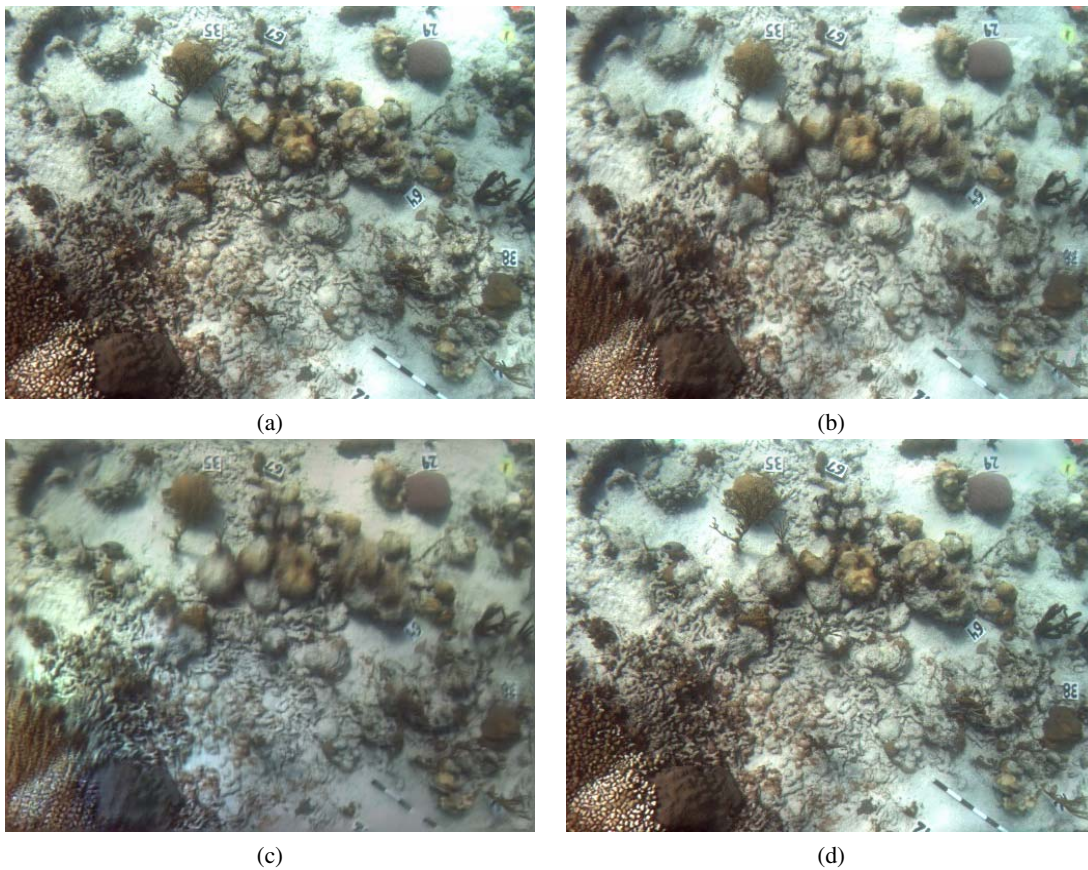

(a)

(b)

(c)

(d)

Fig. 6. Comparision of methods under noticeable registration errors — Original frame (a), temporal median of intensities (b), result of the median of gradients method (c), and result of our method (d).

[5] R. Garcia, T. Nicosevici, and X. Cufi. On the way to solve lighting problems in underwater imaging. In *IEEE OCEANS Conference*, pages 1018–1024, Biloxi, Mississipi, USA, 2002.

[6] K. Zuiderveld. Contrast limited adaptive histogram equalization. *Graphics gems IV*, pages 474–485, 1994.

[7] R. Eustice, O. Pizarro, H. Singh, and J. Howland. UWIT: Underwater image toolbox for optical image processing and mosaicking in Matlab. In *Proceedings of the 2002 International Symposium on Underwater Technology*, pages 141–145, Tokyo, Japan, April 2002.

[8] Y. Rzhanov and Fan Gu. Enhancement of underwater videomosaics for post-processing. In *Oceans 2007*, pages 1–6, October 2007.

[9] Y. Schechner and N. Karpel. Attenuating natural flicker patterns. In *Proc. of OCEANS '04. MTTS/IEEE TECHNO-OCEAN '04*, volume 3, pages 1262–1268 Vol.3, Nov. 2004.

[10] Yasuyuki Matsushita, Ko Nishino, Katsushi Ikeuchi, and Masao Sakauchi. Shadow elimination for robust video surveillance. In *MOTION '02: Proceedings of the Workshop on Motion and Video Computing*, page 15, Washington, DC, USA, 2002. IEEE Computer Society.

[11] Y. Weiss. Deriving intrinsic images from image sequences. *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, 2:68–75 vol.2, 2001.

[12] A. Renyi. *Probability Theory*. North-Holland, Amsterdam, 1970.

[13] N. Gracias and J. Santos-Victor. Underwater video mosaics as visual navigation maps. *Computer Vision and Image Understanding*, 79(1):66–91, July 2000.

[14] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision, 2nd edition*. Cambridge University Press, 2003.

[15] H. Farid. Blind inverse gamma correction. *IEEE Transactions on Image Processing*, 10(10):1428–1433, 2001.

[16] H. Farid and A.C. Popescu. Blind removal of image non–linearities. In *International Conference on Computer Vision*, volume 1, pages 76–81, Vancouver, Canada, 2001.

[17] P. D. Kovesi. MATLAB and Octave functions for computer vision and image processing. School of Computer Science & Software Engineering, The University of Western Australia. Available from: http://www.csse.uwa.edu.au/~pk/research/matlabfns/.

[18] D. Lirman, N. Gracias, B. Gintert, A. Gleason, R. P. Reid, S. Negah-daripour, and P. Kramer. Development and application of a video–mosaic survey technology to document the status of coral reef communities. *Environmental Monitoring and Assessment*, 159:59–73, 2007.

[19] A. Gleason, D. Lirman, D. Williams, N. Gracias, B. Gintert, H. Madjidi, R.P. Reid, G.C. Boynton, S. Negahdaripour, M. Miller, and P. Kramer. Documenting hurricane impacts on coral reefs using two–dimensional video–mosaic technology. *Marine Ecology*, 28(2):254–258, June 2007.

[20] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *ACM Trans. Graph.*, 21(3):249–256, 2002.