

9-2007

# Enhancement of Underwater Video Mosaics for Post-Processing

Yuri Rzhanov

*University of New Hampshire, Durham, Yuri.Rzhanov@unh.edu*

Fan Gu

*University of New Hampshire, Durham*

Follow this and additional works at: <https://scholars.unh.edu/ccom>



Part of the [Oceanography and Atmospheric Sciences and Meteorology Commons](#)

---

## Recommended Citation

Rzhanov, Y.; Fan Gu, "Enhancement of Underwater Videomosaics for Post-processing," in OCEANS 2007, pp.1-6, Sept. 29 2007-Oct. 4 2007 doi: 10.1109/OCEANS.2007.4449276

This Conference Proceeding is brought to you for free and open access by the Center for Coastal and Ocean Mapping at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Center for Coastal and Ocean Mapping by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact [nicole.hentz@unh.edu](mailto:nicole.hentz@unh.edu).

# Enhancement of Underwater Videomosaics for Post-processing

Y. Rzhanov and Fan Gu

Center for Coastal and Ocean Mapping, University of New Hampshire  
24 Colovos Road  
Durham, NH 03824 USA

**Abstract**—Mosaics of seafloor created from still images or video acquired underwater have proved to be useful for construction of maps of forensic and archeological sites, species' abundance estimates, habitat characterization, etc. Images taken by a camera mounted on a stable platform are registered (at first pair-wise and then globally) and assembled in a high resolution visual map of the surveyed area. While this map is usually sufficient for a human orientation and even quantitative measurements, it often contains artifacts that complicate an automatic post-processing (for example, extraction of shapes for organism counting, or segmentation for habitat characterization). The most prominent artifacts are inter-frame seams caused by inhomogeneous artificial illumination, and local feature misalignments due to parallax effects – result of an attempt to represent a 3D world on a 2D map. In this paper we propose two image processing techniques for mosaic quality enhancement – Median Mosaic-based Illumination Correction suppressing appearance of inter-frame seams, and Micro Warping decreasing influence of parallax effects.

## I. INTRODUCTION

The purpose of underwater photomosaicing is to obtain a visually plausible composite image with two desirable properties [1]: 1) The mosaic should not exhibit features that belong only to individual frames but not to the imaged scene, and 2) The mosaic should retain the features of the imaged scene.

This paper proposes two techniques of conditioning of input image sequences such that the blended mosaic satisfies the conditions specified above. Note that the success in image blending may have different meaning depending on the intended user of the product, e.g. a scientist trying to deduce large-scale interrelationships; a computer program extracting shapes according to some specific rule; or a high-school student learning about the deep-sea environment. In any case it is expected that the mosaics obtained with blending techniques appear continuous and homogeneously illuminated.

A major difficulty in processing underwater images is due to specific transmission properties of light in this medium. Light propagation is strongly affected by absorption and scattering. Besides, lack of ambient lighting requires the use of artificial illumination, which is often spatially non-uniform. When camera and source of light move, seafloor features appear differently illuminated on different image frames.

Light attenuation and backscatter limit range of photography such that each image covers only few square meters of the seafloor. To obtain an image of a larger area, hundreds or thousands of individual frames may be needed. Navigation information is not sufficiently accurate to assist in stitching frames together. A composite view can only be obtained by exploiting the redundant multiple overlapping images distributed over the scene. However, non-planarity of the seafloor and short distance between the seafloor and the camera lead to the parallax effects.

Due to these issues, underwater mosaics often contain artifacts that make post-processing difficult or even impossible. Algorithms for automated object recognition and shape extraction can be easily confused by feature doubling and seams arising due to inhomogeneous illumination. However, these algorithms are typically tolerant to scaling and small shape distortions. The purpose of this paper is to describe techniques for creation of composite images without visible seams and ghosting artifacts, at the expense of introducing some minor local distortions.

## II. MEDIAN MOSAIC-BASED ILLUMINATION CORRECTION

Several blending methods have been proposed for decreasing influence of seams, with various weighting functions, in both spatial and frequency domains [2]. Another approach is to pre-process the images before blending, compensating for differences in luminosity [3]. Existing variations of this approach include: 1) Correction of acquired images according to a smoothed image estimated through a set of consecutive frames and by disregarding the shade component in the illumination-reflectance model [4]; 2) Local histogram equalization [5] [6]; 3) Homomorphic filtering [7] (assuming that the illumination factor varies smoothly through the field of view, this method suppresses the low frequencies while keeping the high frequencies); 4) Subtraction of the illumination field by polynomial adjustment [8] (a low-order 2D polynomial spline is subtracted from the acquired image). Other techniques used for image mosaicing are: 1) Estimation of a single high dynamic range radiance map from the differently exposed images [9][10]. 2) Adjustment of exposure and contrast of images by warping their histogram of pixel values [11] [12].

Median mosaics have been often used for processing underwater imagery. The idea is to use for final mosaic the median

value from the stack composed of contributions from all overlapping frames. With the sufficiently dense coverage median mosaic is less affected by outliers than the mosaic created by any weighted averaging. Median mosaic represents salient properties of the seafloor, rather than transient properties characteristic to individual frames. Thus, the photomosaics will appear more homogeneous if low-frequency trend of an individual frame is replaced by a globally consistent trend obtained from a median mosaic.

In this paper, the trends of overlapping frames are matched to some common surface, so they match each other better on a local scale. De-trending is a statistical or mathematical operation of removing a trend from the series. It is often applied to remove a feature thought to be distorting or obscuring the relationships of interest. In [13], the authors compared the methods of linear, quadratic, and cubic de-trending, as well as wavelet and spline de-trending. De-trending of input frames (“flattening”, replacing the trend with a flat horizontal surface, which has been proposed in [8]) is a particular case of the proposed technique. The important difference is that de-trending may eliminate low-frequency salient features of the imaged surface, while median mosaic correction retains these features. It must be noted that smooth vertical variations in camera trajectory may cause gradual brightening or darkening of imagery which is indistinguishable from a true change in seafloor properties. However, with a sufficiently high density coverage (several swaths acquired from a variety of altitudes) this factor becomes insignificant, as median mosaicing ignores extreme values of luminosity.

In our experiments we have found that second order polynomial fitting to find a trend works reasonably well for a variety of illumination sources and is relatively fast. The calculations are based on the logarithm of the input image luminosity which converts the multiplicative operators to additive.

Let  $l_{ij} = \ln L(i, j)$  denote the logarithm of the pixel value at position  $(i, j)$  of some input image. Let (1) define a second-order 2D polynomial fit to the array of  $l_{ij}$ .

$$T_{ij} = P_1 i^2 + P_2 ij + P_3 j^2 + P_4 i + P_5 j + P_6 \quad (1)$$

Best fitting set of coefficients  $\vec{P} = \{P_k, k = 1...6\}$  (in the least squares sense) can be found from the solution of an over-constrained system of equations for  $l_{ij} : \vec{l} = S\vec{P}$ , where  $\vec{l} = (l_{00}, l_{01}, ..., l_{ij}, ..., l_{M-1, N-1})^T$ , and  $S$  is a matrix dependent only on image dimensions  $(M, N)$ . Moore-Penrose pseudo inversion gives  $\vec{P} = (S^T S)^{-1} S^T \vec{l}$ . Obviously the  $(S^T S)^{-1} S^T$  term depends on the order of the trend and image size only, needs to be computed once, and can be used for all frames.

The proposed method is illustrated in Fig. 1 and consists of the following steps:

- 1) Construct a median mosaic from the original video frames.
- 2) Back-project the corresponding images of the frames from the median mosaic.
- 3) Obtain the 2D trends of back-projected frames and original frames respectively based on the surface fitting parameters in log space.
- 4) Replace the 2D trend of the original frame with that of the corresponding back-projected frame.
- 5) Use the trend-corrected image for constructing the final mosaic.

### III. MICRO WARPING

It is known that only if the scene is flat, its images taken by an unrestricted camera can be combined in an ideal mosaic. Any deviation from flatness gives rise to parallax effects, and during a mosaicing process every non-flat feature is projected onto a composite view on two or more locations, causing ghosting artifacts. In order to remove influence of parallax on a mosaic, the authors of [14] proposed a de-ghosting (local alignment) method which computes the flow between all pairs of images using patch-based alignment method and then allows to infer desired local warps from these computations.

Graph cut-based techniques for building a mosaic (see for example, [15]) do not suffer from ghosting, but cause non-flat features to be misaligned across the seams. These artifacts complicate post-processing of composite images such as, for example, shape extraction. We propose the technique for distortion (warping) of original images such that introduced distortions compensate for misalignments caused by parallax. It is assumed that images are already registered (as in MMBIC), and three-dimensional content is not prominent so that misalignments across the seams do not exceed few percent of a typical image dimension.

Warping of an image consists of choosing a set of control points and corresponding shift vectors determining final locations of these control points. Realization of the fact that it is blurring of prominent features that determines mosaic quality the most, suggests choosing the most noticeable features of the frames as control points. Number of these features is typically few dozens per image. Among all the available image warping techniques we have chosen a thin-plate spline algorithm [16]. A Thin-plate Spline (TPS) is an interpolation method that finds a “minimum energy” smooth surface that passes through all given points. TPS with 3 control points is a plane, for more than 3 is generally a curved surface and for less than 3 is undefined.

The proposed algorithm is shown in Fig. 2 (for the sake of illustration on only two frames) with the following steps:

- 1) Points of interest are extracted from the overlapping frames. We have used SIFT-like features similar to ones described in [17] or SURF feature extraction algorithm proposed in [18].

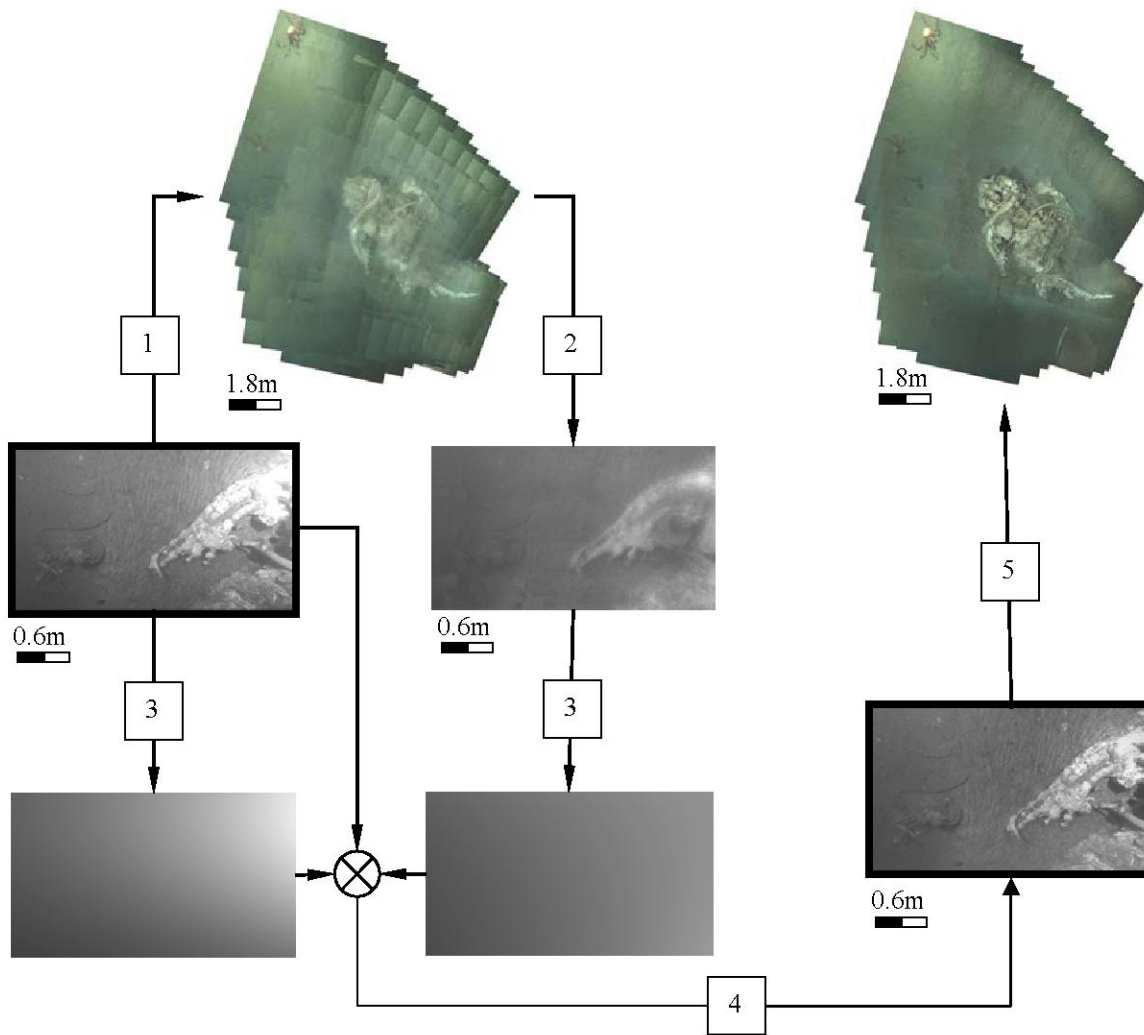


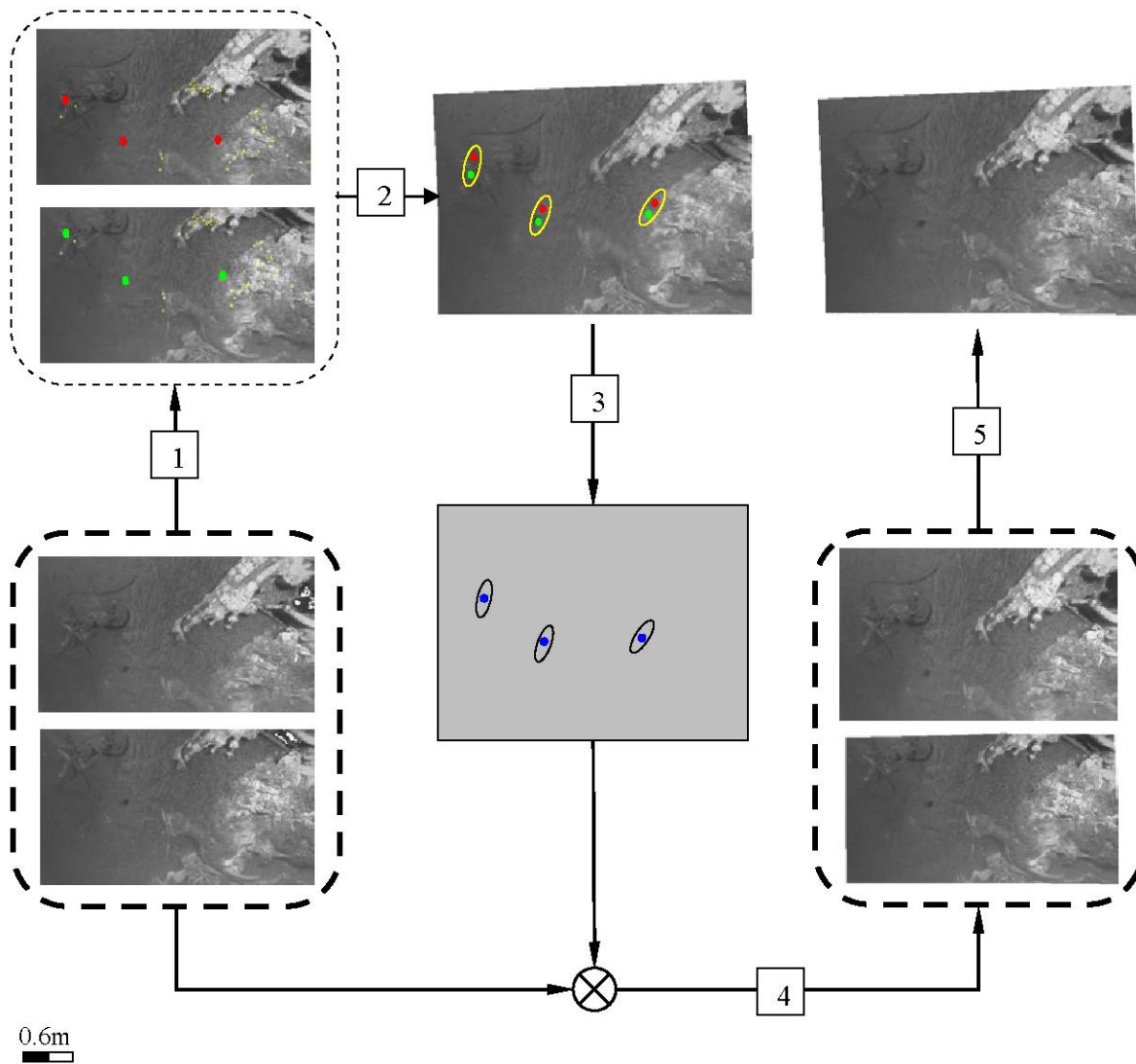
Fig. 1 Framework for MMBIC. Explanations are in the text.

- 2) Two sets of points are matched using local invariant point descriptions with the additional constraint that matching points must not be separated in the mosaic by a distance larger than pre-determined threshold. Note that the matching stage does not require use of any robust procedure like RANSAC.
- 3) For all matched pairs, a new location of control points (features) is placed at a mid-distance between the original locations in the mosaic-based coordinates, these new locations are found in the space of the original images, thus determining shift vectors for all control points.
- 4) Thin-plate spline warping is applied to participating images. The size of images is left unchanged, so some data may disappear, and some pixels may be marked as having no data.
- 5) Warped images are mosaiced again, with the same transformation as before. Features chosen as control points are now mapped onto the same mosaic location.

#### IV. DISCUSSION OF RESULTS

The purpose of MMBIC is to obtain more homogeneous photomosaics with invisible seams. It is based on correction of appearance of single frames prior to mosaicing. One might expect that frames with high degree of overlap must exhibit similar histograms after the correction. Fig. 3 shows histograms of three overlapping frames in a sequence, before and after the application of MMBIC. Histogram representation of an image disregards spatial distribution of pixel values (which explains relative failure of histogram warping techniques), but higher similarity between the histograms after MMBIC indicates that the images have indeed become more homogeneous. Histogram warping methods force histograms of all local areas to be equal, while MMBIC, without direct operations on distributions of pixel values, causes histograms of overlapping images to become





**Fig. 2 Framework for MW.** Explanations in the text. Red and green points represent the features extracted from each frame, and matching points are encircled by ellipses. Blue points are the average of the matching points, and they are identified as the target coordinates for the warping.

more similar, but allows histograms of spatially separated images to be different.

Fig. 4 shows full mosaics of 120 frames, before and after MMBIC. Mosaicing method is graph cut-based.

Experimentation with a number of datasets has shown that after the correction mosaics always appear more visually pleasing, at the expense of two additional operations: construction of a median mosaic and correction of original image frames. A degree of improvement depends mostly on a coverage density. Robust pair-wise registration of images requires 50% pair-wise overlap which guarantees 200% coverage. This density, however, is not sufficient for successful application of MMBIC, and the results in this case are similar to that of simple detrending. Mosaic quality improvement becomes noticeable at 80% sequential overlap, which corresponds to 500% coverage. Fig. 4 (right) clearly shows that the mosaic center (highest density of coverage) appears homogeneous, and the periphery exhibits some artificial seams, however smoothed in comparison with the mosaic from uncorrected footage (left).

The goal of MW is to alleviate ghosting and misalignment artifacts due to parallax issues. Fig. 5 shows the mosaic of two frames from the “Sparrowhawk” sequence. Frames were registered via left plane wing (which has more prominent features), so that the structure on the right which is at the different distance from the camera, remained misaligned. Application of MW has fixed this misalignment.

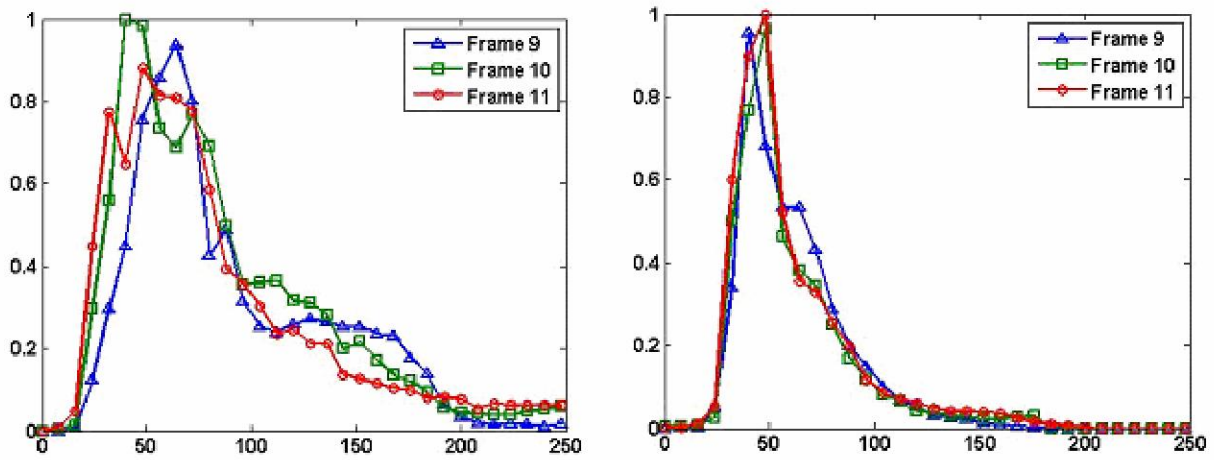


Fig. 3. Normalized histograms of single frames before and after illumination correction. Left: histograms of the original frames. Right: histograms of the corrected frames.

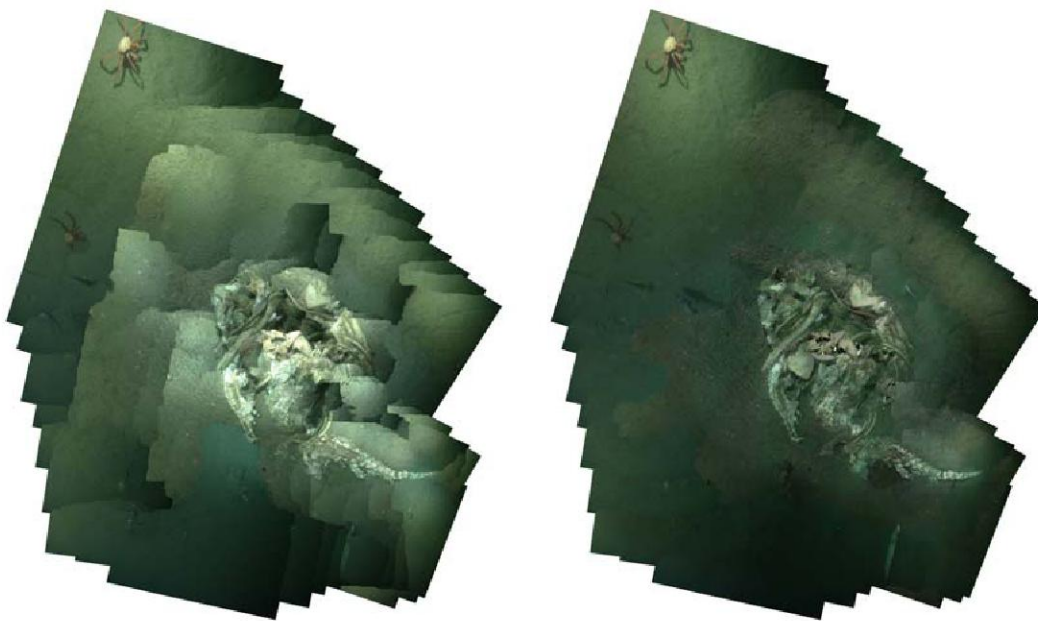


Fig. 4. Photo mosaics ("Whalefall" sequence, 120 frames) constructed using graph-cut technique before (left) and after (right) application of MMBIC.

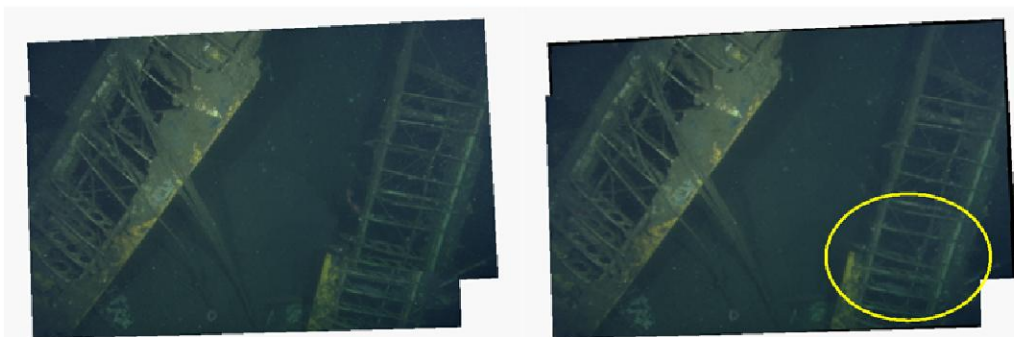


Fig. 5. Photo mosaics ("Sparrowhawk" sequence, frames 57-58) showing results of Micro Warping. Left: before, right: after. The area with the most noticeable improvement is surrounded with yellow ellipse.

The results of MW show that the ghosting/misalignment artifacts are being reduced, so that composite images appear more consistent and allow for extraction of more accurate information. Due to limitations of the feature extraction techniques and specifics of thin-plate spline warping, this method performs well for small misalignments only (up to 15% of characteristic image dimension) and for the features appearing in only 2-3 consecutive frames. A non-planar feature appearing at the top of one frame and at the bottom of another usually changes in appearance so much that it does not make sense to warp both frames in an attempt to bring the instances of this feature together in the mosaic space. This, however, does not pose a problem for typical density of coverage exceeding 200-300 percent. Only features separated by distances less than 15% of a frame dimension were chosen as control points for warping. This improved alignment of elongated features across the seams in a graph cut-based blending process and did not cause unrealistically strong distortions of images.

## CONCLUSIONS

We have presented two techniques for improvement of underwater mosaic quality that make mosaics more suitable for post-processing such as pattern recognition and shape extraction. First technique, Median Mosaic-based Illumination Correction, allows for diminishing of inter-frame seams appearance, and the second, Micro Warping, compensates for artifacts related to parallax effects.

## ACKNOWLEDGEMENTS

The authors thank Dr. R. Vrijenhoek (MBARI) and B. Terrell (NOAA's National Marine Sanctuary program) for permission to use the image data. This work was funded by NOAA grant NA97OG0241.

## REFERENCES

- [1] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless image stitching in the gradient domain", *In Proc. Eighth European Conference on Computer Vision*, vol. 4, pp. 377-389, 2004.
- [2] P.J. Burt, and E.H. Adelson, "A multi-resolution spline with application to image mosaics", *ACM Trans. On Graphics*, vol. 2(4), pp. 217-236, 1983.
- [3] R. Garcia, T. Nicosevichi, and X. Cufi, "On the way to solve lighting problems in underwater imaging", *Proc. Oceans'02 MTS/IEEE*, pp. 1018-1024, 2002.
- [4] R.C. Gonzalez, R.E. Woods, *Digital Image processing*, Prentice Hall, 2002.
- [5] H. Singh, J. Howland, D. Yoerger, and L. Whitcomb, "Quantitative photomosaicing of underwater imaging", *Proc. Oceans'98 MTS/IEEE*, v. 1, pp. 263-266, 1998.
- [6] R. Eustice, H. Singh, and J. Howland, "Image registration underwater for fluid flow measurements and mosaicking", *Proc. Oceans'00 MTS/IEEE*, v. 3, pp.1529-1534, 2000.
- [7] A. V. Oppenheim, R. W. Schaffer, and T. G. Stockham Jr., "Non-linear filtering of multiplied and convolved signals". *Proc. IEEE*, vol. 56, no. 8, pp. 1264-1291, 1968.
- [8] Y. Rzhanov, L. Linnett, and R. Forbes, "Underwater video mosaicing for seabed mapping", *Proc. of the IEEE Conf. on Image Processing*, vol. 1, pp. 224-227, 2000.
- [9] R. S. Mann and R. W. Picard, "Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures", *Proc. of IS&T 46th Annual Conference*, pp. 422-428, 1995.
- [10] P. E. Debevec, and J. Malik, "Recovering high dynamic range radiance maps from photographs", *Proc. SIGGRAPH 97*, pp. 369-378, 1997.
- [11] I. Cox, S. Roy, and S. Hingorani, "Dynamic histogram warping of image pairs for constant image brightness". *Proc. of the IEEE International Improving Feature Tracking with Robust Statistics Conference on Image Processing*, vol. 2, pp. 366-369, 1995.
- [12] R. M. Grundland and N. A. Dodgson, "Automatic Contrast Enhancement by Histogram Warping", *Proc. International Conference on Computer Vision and Graphics*, v. 32, pp. 293-300, 2004.
- [13] R. J. Tregellas, R. Freedman, J. Tanabe, D. Miller and F. G. Meyer. "Comparison of detrending methods for optimal FMRI preprocessing", *NeuroImage*, pp. 902-907, 2002.
- [14] H.-Y. Shum and R. Szeliski, "Construction of panoramic mosaics with global and local alignment", *International Journal of Computer Vision*, v. 35(2), pp. 101-130, 2000.
- [15] A. Agarwala, et al., "Interactive digital photomontage", *Proc. SIGGRAPH'04*, pp.294-302, 2004.
- [16] G. Wahba, "How to smooth curves and surfaces with splines and cross-validation", *Proc of 24th Conference on Design of Experiments*, pp. 167-192, US Army Research Office, 1979.
- [17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, v. 60(2), pp. 91-110, 2004.
- [18] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," *Proc. ECCV*, vol. 1, pp. 404-417, 2006.