A Novel Blending Technique for Two-Dimensional Forward-Looking Sonar Mosaicing

Natàlia Hurtós, Xevi Cufí and Joaquim Salvi
University of Girona, Edifici Politecnica IV, Campus Montilivi 17071 Girona, Spain
{nhurtos,xcuf,qsalvi}@eia.udg.edu

***Abstract*—High-resolution forward-looking sonars are becoming a tool of choice for exploring underwater environments under low visibility conditions. Their imagery can be mosaiced to obtain a global overview of submerged areas of interest and the spatial arrangement of different target features. However, in order to achieve an informative and smooth image composition, the individual sonar frames must be fused. Unlike the blending in optical mosaics, this implies dealing with a high number of overlapping images as well as with sonar specific artifacts arising from its image formation geometry. This work presents a novel blending pipeline designed to cope with these artifacts involving strategies to diminish the impact of all the photometric irregularities that might be present when mosaicing forward-looking sonar imagery. Results of blended mosaics, including data gathered with different sonar models and presenting several artifacts, are presented here to show the applicability of the method.**

I. Introduction

The increasing development of two-dimensional forward-looking sonars (FLS) which deliver high-resolution acoustic images at near-video frame rate is playing a key role in underwater inspections where water visibility does not allow the use of optical cameras. Inspection of harbour underwater structures, ship hulls, dams or the monitoring of rivers and lakes are some of the applications that can benefit from this growing technology. Following this line, several authors [1], [2], [3] have studied the development of mosaicing techniques specifically suited to FLS imagery with the aim of providing an overall view of an area of interest even in the presence of turbid waters.

The general pipeline for mosaicing consists of several steps. First, the pairwise registration of sonar images is performed to obtain an initial guess of the trajectory, either by using feature-based techniques [1] or frequency-based registration [3]. This trajectory can then be refined through global alignment techniques by using information of loop closure situations [3]. As a result, the acquired images can be projected and rendered onto a single and common reference frame. However, without any image fusion mechanism, the seams along the different images boundaries become noticeable due to photometrical differences between the individual sonar frames or due to geometrical registration inaccuracies. Therefore, it is necessary to perform one last step to give a continuous and uniform appearance in the form of a single large mosaic. This is achieved by means of image blending techniques. It is worth highlighting that generating the sonar mosaic with a convincing and natural appearance has not only aesthetic but informative purposes. The interpretation of a given scene becomes more intuitive and effective when its features are emphasized and it has a global smooth and continuous appearance.

The basic principles of image blending were established four decades ago [4] and the topic has been extensively studied in the field of optical imaging, including underwater environments [5]. However, the inherent nature of FLS imagery poses some particular challenges that need to be specifically addressed to obtain a proper sonar blended mosaic. In this paper we present a novel methodology to blend mosaics obtained from FLS imagery. We start in Section II by reviewing related work and analyzing the specific problems to be faced in the blending of FLS mosaics compared to the traditional approaches adopted on optical photomosaics. Section III presents the proposed methodology, describing the main steps performed to correct photometric irregularities both at frame level and at a global mosaic scale. Results of mosaics rendered by using the proposed blending pipeline are show in section IV, involving data gathered with different sonar models and affected by different artifacts. Finally, section V provides some concluding remarks and points out future work.

II. Background

There is a wide variety of image blending techniques in the literature, but at a high level two main approaches can be distinguished [5]. On one hand, we have transition smoothing methods (also known as feathering or alpha blending methods) which attempt to minimize the visibility of the image boundaries by smoothing the common overlapping region of the stitched images. On the other hand, there are optimal seam finding methods which attempt to find the optimal location to place a cut along the two images so that it minimizes the photometrical and geometrical changes between them. Furthermore, there are also hybrid techniques which take advantage of the benefits of each approach.

Hence, regardless of the particular technique, one can see that optical blending generally deals with a low number of images at a given position (most of the times pairwise) and treats only their intersecting region. This prevents us from directly leveraging traditional blending techniques designed for video images since blending a FLS mosaic requires dealing with multiple overlapping images involving high overlap percentages. High overlap is usual in FLS data, not only because of the high frame rate of the FLS sensors, but also because when acquiring images in an across-range fashion high overlap is a must to achieve good coverage due to the sonar fan-shaped footprint. Moreover, presuming that a correct registration has been performed, it is of interest to keep as much of overlapping images as possible to be able to improve the signal-to-noise ratio (SNR) of the final mosaic. This is again opposed to other approaches typically adopted on optical mosaicing such as trying to select only the best image for a given location. Therefore, for blending FLS mosaics it is necessary to deal not only with the seam areas, but with the whole image content.

In addition to this main divergence on the mosaicing approach there are several sonar-specific issues at frame level that can also have an impact on the blending step. FLS imagery is affected by two sorts of illumination artifacts. Firstly, sonar frames often present a constant inhomogeneous insonification pattern due to the different sensitivity of the lens or transducers across the field of view. Additionally, FLS images are subject to illumination alterations due to changes in the point of view or changes in the underlying scene topology. When an object or a scene is imaged while the sonar is in motion, the object's vertical displacement fluctuates within the elevation angle of the sonar, receiving more or less incidence and thus causing variations on the illumination profile, similar to the well-known parallax effect that occurs in optical images.

Furthermore, the main parameters that configure the imaging geometry, namely tilt angle of the sensor, altitude from the seafloor and minimum and maximum range of the image, play a key role for acquiring a proper image. FLS insonify the scene with an acoustic wave spanning its field of view in azimuth (θ) and elevation (ϕ) directions (Fig. 1). The acoustic return is sampled by an array of transducers as a function of range and bearing, resulting in an image with the backscattered intensities at each point (r, θ). Usually due to the narrow elevation angle of the FLS sensors (around 12 to 20 degrees) the sonar is tilted at a grazing angle from the scene so as to maximize the coverage of the insonification area. However, due to inappropriate setup of the imaging configuration (i.e. navigating at too high altitudes or not tilting the sonar enough for the established ranges) or due to the deviation of the scene from the planar assumption that the mosaicing techniques presume [3], there are cases where the imaged area becomes just a portion of the sonar image footprint (Fig. 2). Although this is a problem that could be avoided by adopting a proper setup according to the underlying scene, our experience suggests that this is sometimes difficult, especially when the inspection area is unknown and there is no mechanism to dynamically detect and adapt to the underlying topology (such as a pan and tilt unit). Hence, it is important for a blending technique to be able to cope with these artifacts so as to preserve the area of the image where there is information and prevent blind areas to cover the real content when all images are registered.

Finally, FLS images suffer also from heterogeneous resolution. This is due to the fact that the sonar image is originally formed in the azimuth-range sampling space. Therefore, when it is mapped to a Cartesian coordinate system, the resolution is reduced depending on the distance of the pixel from the sonar origin (i.e., one pixel is mapped to a group of pixels). Although this does not usually have a strong visual impact on the mosaic, it is also a particularity that must be considered in the blending process.

The state of the art does not include precise solutions to cope with all the aforementioned factors and, in fact, little work can be found in the literature regarding sonar image blending. In [6] side-scan sonar data is mosaiced and blended using a wavelet-based technique that allows to select which

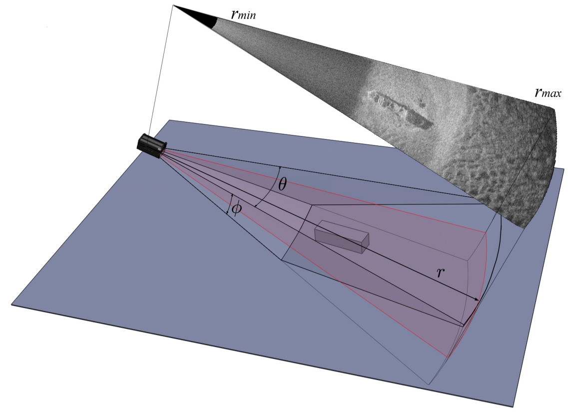


Fig. 1. Imaging sonar geometry (r: range, θ: azimuth, ϕ: elevation).

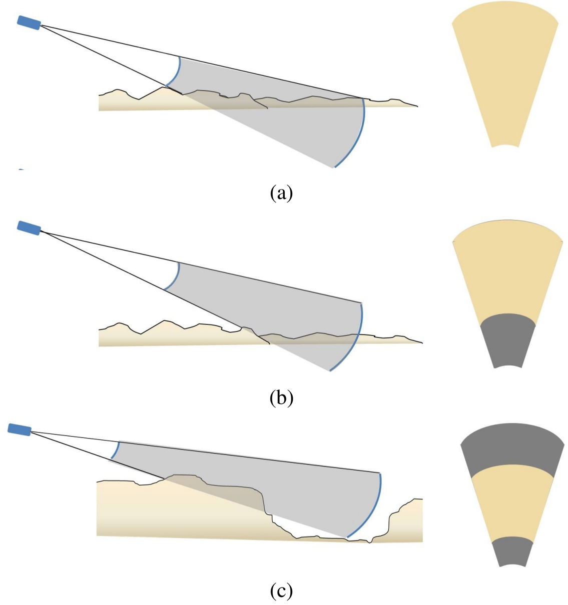


Fig. 2. Schemes illustrating different sonar-scene configurations and their corresponding acquired images showing blind regions in dark grey. (a) Proper imaging configuration where the altitude over the scene, the tilt angle and the minimum and maximum ranges are adjusted for a good image coverage. (b) Imaging configuration where the sonar is located too high or not tilted enough for the established ranges. (c) A change in the scene relief causing the apparition of a blind region.

kind of features are emphasized in the final mosaic. Kim et al. [7] proposed a probabilistic approach in the context of a superresolution technique for FLS frames. They model the blending problem of fusing a low-resolution image into a high-resolution one in terms of a conditional distribution with constraints imposed by the illumination profile of the observed frames so as to maximize the SNR of the resulting image. In our previous work on FLS mosaicing [3] as well as in the work of Wei Yong [8], results have been rendered by averaging the intensities of all overlapping sonar frames at every mosaic pixel. Averaging the overlapping sonar intensities results in denoising of the final mosaic, achieving an improvement in terms of SNR compared to a single image frame. Although this approach is a good starting point and may give satisfactory results in some imaging configurations, it diminishes details in those places where there are a large number of frame contributions, as well as shows image boundaries where the number of overlapping images is not constant. Therefore, a better blending procedure should be devised in those situations in order to achieve a visually pleasant result.

III. METHODOLOGY

The proposed methodology consists of a set of strategies to address the problems explained in the previous section associated with FLS image composition. Our approach takes as a basis the fusion by intensity averaging and incorporates strategies to correct for the different artifacts and modify the number and/or the intensity of the averaged pixels for the final image composition.

It is important to stress that the proposed blending concentrates on solving the different photometric artifacts that can arise both at individual frame level and at global mosaic scale but does not focus in possible problems caused by geometric registration errors. Therefore, a correct registration is assumed from previous steps, otherwise the averaging principle leads to blurred areas of mixed content. In the same way, other geometric issues such as object shadow alterations due to the viewpoint change are handled implicitly by the averaging principle. For instance, when imaging a protruding object while navigating over it, its shadow gets shortened as the sonar becomes closer to the object. The final shadow representation in the mosaic will then be the mean of all shadow positions yielding an intermediate solution which we consider to be a reasonable description of the scene.

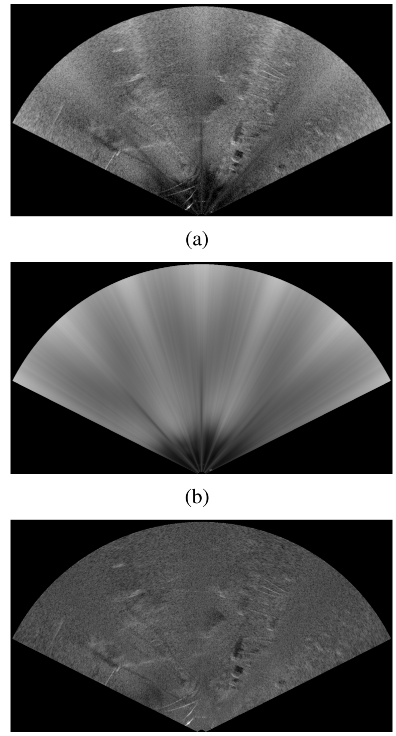
Each of the strategies presented here can be enabled or disabled in the blending pipeline according to the characteristics of the dataset. Therefore a dataset gathered in ideal conditions (i.e. with a sonar that would not present inhomogeneous insonification patterns, with the proper altitude, tilt and range settings, imaging a planar scene and performing just a single trackline at constant speed in order to keep a uniform number of overlapped images) would be blended through a standard intensity averaging only benefiting from the local contrast enhancement step to emphasize its features.

It is worth noting that the proposed blending pipeline is designed to work in an off-line fashion as it requires using all gathered frames with the aim of producing a final high-quality map of the inspected area. Note also that although the images can be projected and fused using the registration result at pixel level, if the mosacing method provides subpixel accuracy the final image locations can be rendered in a higher resolution grid thus being able to take into account these subpixel displacements and obtain a higher resolution mosaic.

A. Individual image pre-processing

The described photometric artifacts that occur at image level can affect the global appearance of the mosaic composition. Hence, it is important to pre-process individual sonar frames to correct for some possible irregularities.

1) Inhomogeneous insonfication pattern correction: Some sonar models show evidence of non-uniform insonification patterns due to the different sensitivity of the transducers across the field of view (Fig. 3(a)). If a sufficient number of images is available, the underlying illumination profile can be computed



(a)

(b)

(c)

Fig. 3. Correction of non-inhomogeneous insonification. (a) Original frame. (b) Estimated illumination pattern. (c) Corrected frame.

by averaging all the dataset frames (Fig. 3(b)). Then, the illumination profile can be compensated in the original image thus yielding a pattern-free image (Fig. 3(c)). If the pattern is strong, this step should be performed earlier in the mosaicing pipeline (i.e. prior to the pairwise registration of the sonar images) since its presence may influence the registration result.

2) Contrast Limited Adaptive Histogram Equalization (CLAHE): Besides non-uniform insonification related to the sensor's hardware, FLS data can exhibit other non-constant illumination patterns. Due to imaging configuration and/or terrain curvature the images can exhibit weaker backscattered intensities in some areas (e.g. weaker intensities further away from the sonar origin). This results in considerable intensity offsets when registering images that insonify the same portion of the scene but from different locations (Fig. 4(a),4(b)), turning into visible seams when blending the mosaic.

To deal with this, we first equalize the intensity histograms of the sonar frames so as to match a uniform distribution, thus minimizing the intensity offsets on the registered areas. To this end, we employ the CLAHE technique [9] whose advantages are twofold: first, it equalizes the images limiting the noise in the areas that are more homogeneous by setting a clip limit on the histogram equalization; second, it locally enhances the contrast of the images alleviating the attenuation of target features due to the low SNR that characterizes FLS images. Although, as stated before, the SNR is greatly enhanced by the averaging nature of the blending, a local contrast enhancement can help to further emphasize the scene features.

Note that this procedure does not preserve the true reflectance values of the scene. However, we believe that for

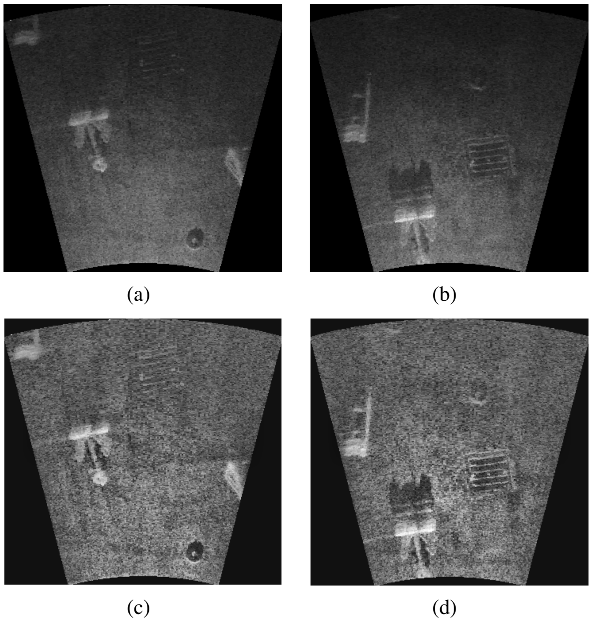


Fig. 4. Image equalisation and contrast enhancement. (a)-(b) Example of two frames imaging the same area from two different viewpoints. Notice the difference in intensities around the grid in the center. (c)-(d) Same frames preprocessed with CLAHE. The images present a uniform distribution that allows to merge them without arising visual seams. Notice also that the local contrast is preserved, emphasizing the scene features.

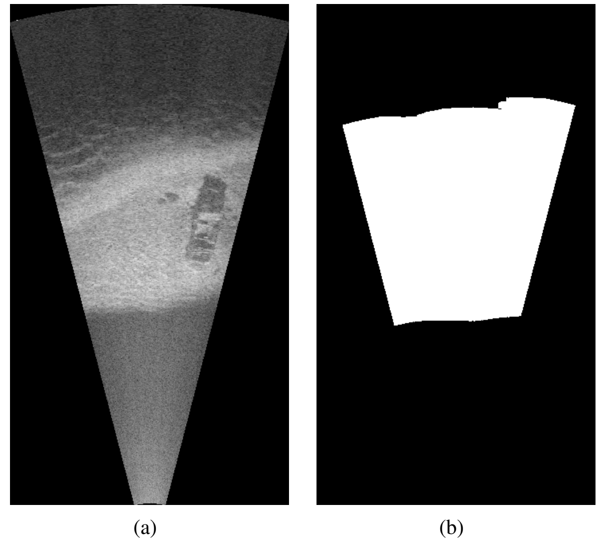


Fig. 5. Example of mask computed to discard blind regions. (a) Original image. (b) Mask (black areas will not be taken into account for the blending).

inspection purposes, it is more important to obtain a continuous and smooth representation that emphasizes the features and facilitates a better scene interpretation rather than preserving the true scene reflectances.

3) Masking out blind regions: Extreme cases of non-uniform intensities across the images are those situations described in Section II, where either an inappropriate imaging configuration or significant relief variations cause blind regions in the sonar frames. Even applying CLAHE and working with the equalized images, those blind regions have a negative impact on the final blending. Since they do not contain information at all, they cause the actual scene content to fade out when they are averaged with other images.

Our strategy for those cases is to compute a saliency mask for each frame (Fig. 5), which will be used to mask out the blind regions when performing the fusion by averaging. The mask M is obtained by applying standard deviation to local neighbourhoods:

$$M(u,v) = \sqrt{\frac{\sum(I(x,y) - \bar{I}(x,y))^2}{n-1}}, \tag{1}$$

where n is the number of pixels in the neighbourhood, $I(x,y)$ is the intensity of the pixel under consideration and $\bar{I}(x,y)$ is the mean of all neighbourhood pixels. The shape and size of the local neighbourhood are parameters that can be adjusted so as to take into account the standard deviation generated by the residual noise of the images. This standard deviation filter acts as a texture classifier. The blind regions of the image, which are characterized by the lack of backscattered intensities report low values. On the other hand, scene backscattered intensities generate higher filter responses even in homogeneous parts. Hence, a threshold is set to segment both type of regions and finally morphological operations are applied to ensure that isolated pixels are not remaining as part of the mask.

B. *Global mosaic blending*

Despite the illumination corrections performed at individual frame level, the fusion of images from different tracklines will unavoidably create noticeable seams along the tracks due to the presence of a higher number of image contributions in the overlapping area (Figure 6(a)). If the tracks are combined along-range, seams may be also noticeable as a consequence of merging two different image quality areas (low/high resolution).

To reduce these artifacts, we compute an overlap map that reflects the number of images projected at each pixel location, taking into account the possible masks that might have been computed previously if the images contain blind regions. In the presence of multiple tracklines, the intersection area will present a significantly higher number of overlapped images compared to their surroundings (Fig. 6(b)).

To avoid these artifacts we propose a mechanism consisting of three main steps: clipping, smoothing and selection. First, the number of overlapping images are clipped to a threshold thus reducing the range of possible different overlaps. While it helps to reduce the overload of pixel contributions at a given location it is also of interest to keep a significant number of overlapped images to diminish the noise of the final mosaic. A trade off solution consists of cutting up to the mean of the overlap map. Second, the new overlap map is smoothed with a gaussian kernel to avoid sharp transitions caused by a different number of pixel contributions. A normalization is required so as to avoid any new computed overlap to exceed the number of actual overlapping images. Finally the mosaic is blended by averaging the number of pixels indicated by the new overlap map. To select from the images that are projected to a given pixel which ones will be discarded and which ones will be taken into account on the final averaging, the following procedure is implemented: for each pixel of the final mosaic we store the list of all values that are projected to that location together with a weight that reflects its position in its original frame. A weighting mask is used to reflect the location, and therefore the resolution, of each pixel (from higher to lower as measurement sparseness increases with the range on Cartesian space). In this way, candidate pixels are sorted according to their weight and the first N ones of higher weight (being N the number of overlapping pixels in the newly computed overlap map) are used to compute the final pixel intensity by averaging. In this way we give priority to those frames that depict the region with higher resolution.

IV. Results

In this section we present several results to validate the effectiveness of the proposed blending pipeline.

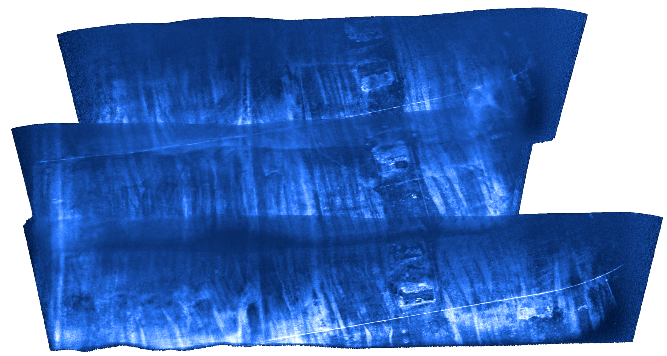
The first example consists of a small dataset gathered with the ARIS sonar [10] navigating over an underwater target lying on a sandy bottom. As the viewpoint changes, different scene features can be observed in the sonar frames, including the sand ripples at the bottom part, the target itself and the sand ripples located at the top, while several blind regions appear as a consequence of an inadequate imaging setup. After mosaicing the frames and blending them by using a straight-forward intensity averaging it can be observed that some of the features disappear and others attenuate as a consequence of being merged with the blind regions (Fig. 7(c)). Processing the mosaic with our blending pipeline with CLAHE and masking of blind regions we obtain the result of Fig. 7(d), in which all the scene features are clearly preserved. In addition to this, note that the SNR has been greatly improved in comparison to an individual sonar frame.

A second test consists of a DIDSON [11] dataset imaging a ship hull along three different horizontal tracklines. Due to the imaging setup involving a small tilt angle (about 15 degrees with respect to the imaged plane) and a slight curvature of the hull the backscattered intensities at longer ranges appear attenuated (Fig. 6(a)). Therefore it is required to enable the CLAHE step in addition to the overlap clipping-smoothing-selection mechanism to correctly blend the different tracklines. The final mosaic blended according to the overlap map of Fig. 6(c) and averaging the values of the best available images in terms of resolution can be seen in Fig. 6(d).

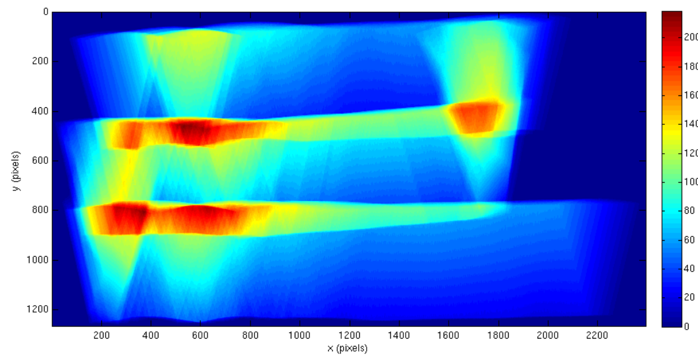
Finally, an example is shown with a BlueView P900-130 [12] performing a trajectory on a harbour environment with multiple tracklines. The images have been preprocessed to correct for a strong inhomogenous insonification pattern before computing the registrations (Fig. 3) and have been equalized and with the CLAHE step. The original overlap map can be seen in Fig. 8(b) while the final blended mosaic is shown in Fig. 8(c). It can be seen that scene features are emphasized and no seams are noticeable despite the presence of multiple tracks and areas of non-constant overlap concentrated in the rotation regions.

V. Conclusion

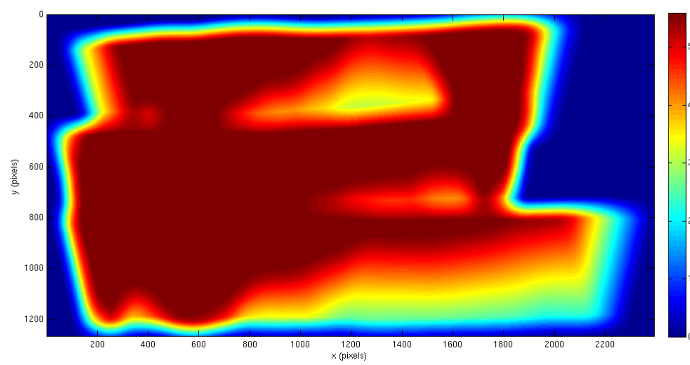
This paper has presented a novel strategy to blend FLS images in order to achieve consistent and visually pleasant



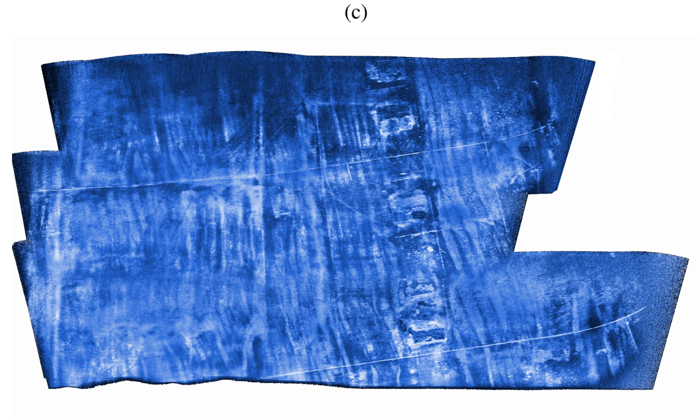
(a)



(b) Estimated illumination pattern.



(c)



(d)

Fig. 6. (a) Detail of a DIDSON sonar mosaic presenting three different straight tracklines. Note the visible seams at the regions of track intersection. (b) Overlap map showing large differences in the number of overlapping images across the mosaic. (b) Computed overlap map clipping the highest overlap values and applying smoothing to avoid abrupt changes. (d) Blended mosaic.

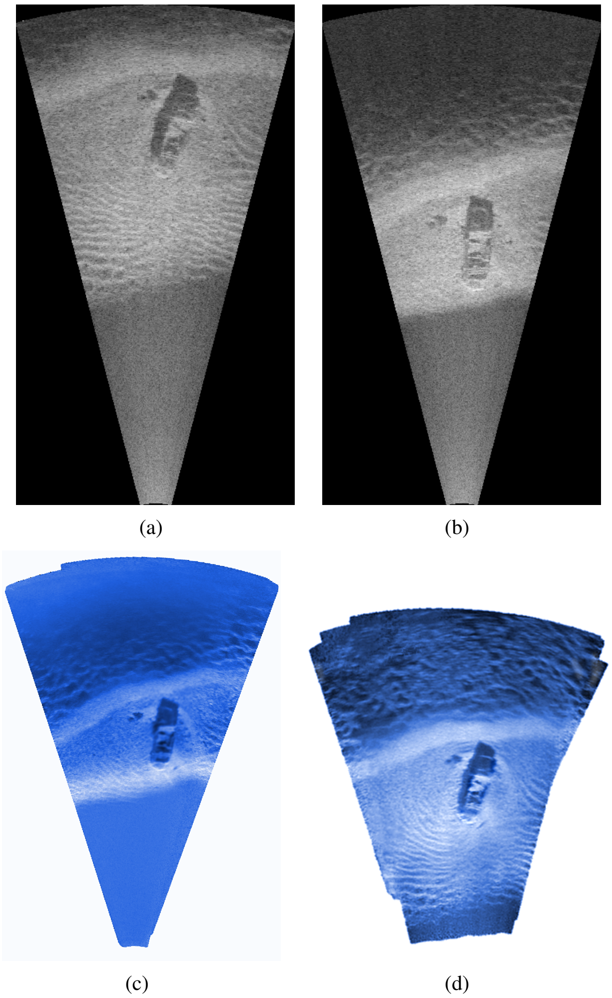


Fig. 7. Example of blending under illumination. (a)-(b) Example of dataset frames affected by varying illumination depending on the imaging viewpoint. (c) Result after standard intensity averaging. (d) Result blended by applying the CLAHE enhancement and the masking of blind regions. Note that features are emphasized, and the sand ripples above and below the center target can be clearly appreciated.

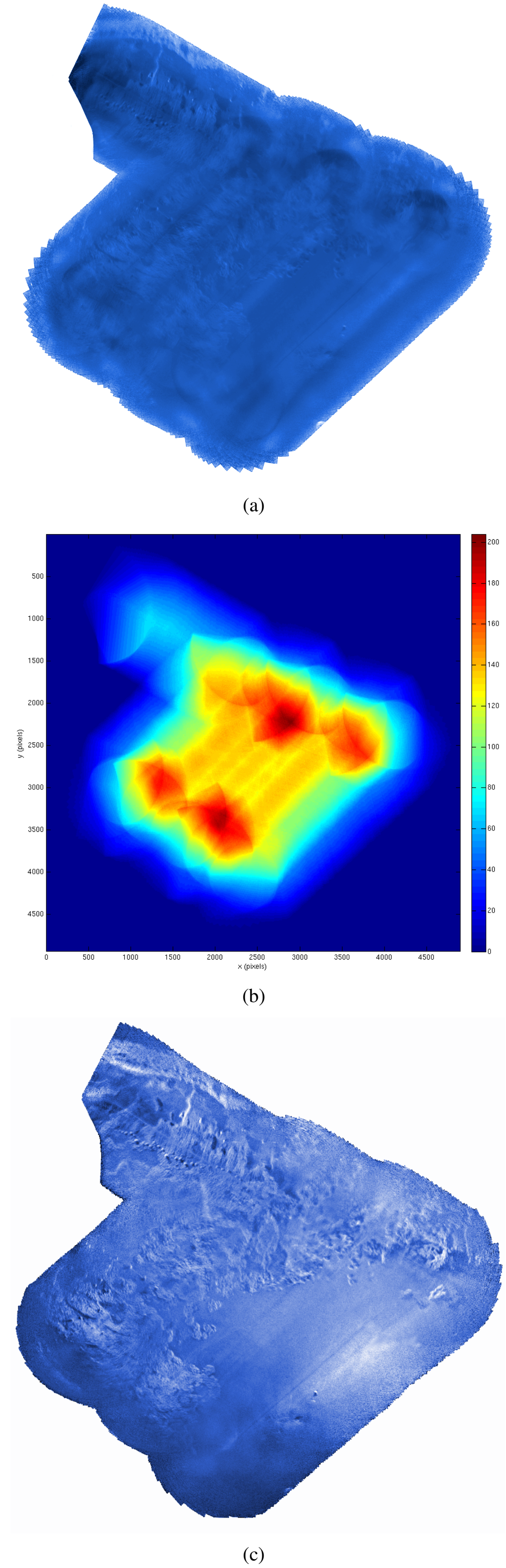


Fig. 8. Mosaic of a harbour area gathered with BlueView P900-130 (a) Mosaic rendered averaging the non-preprocessed images. (b) Original overlap map, note the different level of overlaps caused by rotational movements between tracks. (c) Blended mosaic after insonification pattern correction, CLAHE and clipping-smoothing-selection on the overlap map.

acoustic mosaics with applications to underwater inspection in turbid waters. The main differences with respect to the blending of optical mosaics, such as dealing with a high number of overlapping frames at a given location and the illumination artifacts particular to the imaging geometry of FLS, have been addressed.

The proposed blending pipeline is designed as a set of multiple strategies that can be enabled depending on the different photometrical irregularities encountered in the data to process. In this way, it allows the enhanced rendering of a wider number of situations, involving data gathered with non-ideal imaging configurations, large areas composed of multiple tracklines or vehicle trajectories at non-uniform speeds.

As a future work, a more optimized implementation should be devised for large mosaics where the global blending step can become memory demanding.

Acknowledgment

This work has been supported by the FP7-ICT-2011-7 project PANDORA-Persistent Autonomy through Learning, Adaptation, Observation and Re-planning (Ref. 288273) funded by the European Commission and the Spanish Project ANDREA/RAIMON (Ref. CTM2011-29691-C02-02) funded by the Ministry of Science and Innovation. The authors would like to thank Sound Metrics Corp. and the Center of Maritime Research and Experimentation for providing sonar data.

References

[1] K. Kim, N. Neretti, and N. Intrator, "Mosaicing of acoustic camera images," *IEEE Proceedings Radar, Sonar and Navigation*, vol. 152, no. 4, pp. 263–270, 2005.

[2] S. Negahdaripour, P. Firoozfam, and P. Sabzmeydani, "On processing and registration of forward-scan acoustic video imagery," in *Proc. 2nd Canadian Conf. Computer and Robot Vision*, 2005, pp. 452–459.

[3] N. Hurtos, X. Cuf', Y. Petillot, and J. Salvi, "Fourier-based registrations for two-dimensional forward-looking sonar image mosaicing," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, oct. 2012, pp. 5298 –5305.

[4] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics (TOG)*, vol. 2, no. 4, pp. 217–236, 1983.

[5] R. Prados, R. Garcia, N. Gracias, J. Escartín, and L. Neumann, "A novel blending technique for underwater gigamosaicing," *IEEE JOURNAL OF OCEANIC ENGINEERING*, vol. 37, no. 4, 2012.

[6] E. Coiras, I. T. Ruiz, Y. Petillot, and D. Lane, "Fusion of multiple side-scan sonar views," in *OCEANS'04. MTTS/IEEE TECHNO-OCEAN'04*, vol. 4. IEEE, 2004, pp. 2036–2054.

[7] K. Kim, N. Neretti, and N. Intrator, "Map fusion method for super-resolution of images with locally varying pixel quality," *International Journal of Imaging Systems and Technology*, vol. 18, no. 4, pp. 242–250, 2008.

[8] E. W. Yong, "Investigation of mosaicing techniques for forward looking sonar," Master's thesis, Heriot-Watt University, 2011.

[9] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[10] "Aris," 2013, retrieved July 26, 2013, from http://www.soundmetrics.com/Products/ARIS-Sonars/ARIS-Explorer-3000.

[11] "Didson," 2013, retrieved July 26, 2013, from http://www.soundmetrics.com/Products/DIDSON-Sonars.

[12] "Blueview," 2013, retrieved July 26, 2013, from http://www.blueview.com.