SCIENCE AND TECHNOLOGY ORGANIZATION
CENTRE FOR MARITIME RESEARCH AND EXPERIMENTATION

Reprint Series **CMRE-PR-2019-142**

Forward looking sonar mosaicing for mine countermeasures

Fausto Ferreira, Vladimir Djapic, Michele Micheli,
Massimo Caccia

June 2019

Originally published in:

Annual Reviews in Control, volume 40, 2015, pp. 212-226,
doi: doi.org/10.1016/j.arcontrol.2015.09.014

About CMRE

The Centre for Maritime Research and Experimentation (CMRE) is a world-class NATO scientific research and experimentation facility located in La Spezia, Italy.

The CMRE was established by the North Atlantic Council on 1 July 2012 as part of the NATO Science & Technology Organization. The CMRE and its predecessors have served NATO for over 50 years as the SACLANT Anti-Submarine Warfare Centre, SACLANT Undersea Research Centre, NATO Undersea Research Centre (NURC) and now as part of the Science & Technology Organization.

CMRE conducts state-of-the-art scientific research and experimentation ranging from concept development to prototype demonstration in an operational environment and has produced leaders in ocean science, modelling and simulation, acoustics and other disciplines, as well as producing critical results and understanding that have been built into the operational concepts of NATO and the nations.

CMRE conducts hands-on scientific and engineering research for the direct benefit of its NATO Customers. It operates two research vessels that enable science and technology solutions to be explored and exploited at sea. The largest of these vessels, the NRV Alliance, is a global class vessel that is acoustically extremely quiet.

CMRE is a leading example of enabling nations to work more effectively and efficiently together by prioritizing national needs, focusing on research and technology challenges, both in and out of the maritime environment, through the collective Power of its world-class scientists, engineers, and specialized laboratories in collaboration with the many partners in and out of the scientific domain.

NOTE: The CMRE Reprint series reprints papers and articles published by CMRE authors in the open literature as an effort to widely disseminate CMRE products. Users are encouraged to cite the original article where possible.

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Forward looking sonar mosaicing for Mine Countermeasures☆

Fausto Ferreira[a,*], Vladimir Djapic[b], Michele Micheli[a], Massimo Caccia[c]

[a] NATO Science & Technology Organization Centre for Maritime Research and Experimentation, Viale San Bartolomeo 400, La Spezia 19126, Italy
[b] SPAWARSYSCEN Pacific, 53560 Hull Street, San Diego, CA 92152-5001, USA
[c] National Research Council of Italy, Institute of Intelligent Systems for Automation Electrical CNR-ISSIA, Via Amendola 122/D, Bari 70126, Italy

ARTICLE INFO

Article history:
Received 14 May 2015
Accepted 27 August 2015
Available online 6 November 2015

Keywords:
Mosaicing
Forward looking sonar
Automatic Target Recognition
Mine Countermeasures

ABSTRACT

Forward looking sonars (FLS) are nowadays popular for many different applications. In particular, they can be used for Automatic Target Recognition (ATR) in the context of Mine Countermeasures. Currently, ATR techniques are applied to raw data which generates many false positives and the need for human supervision. Mosaicing FLS data increases target contrast and thus reduces false positive rate. Moreover, it implies a considerable data size reduction which is important if one thinks of exchange of data in real time through an acoustic channel with very limited bandwidth. Results of applying a real-time mosaicing algorithm to FLS data generated during Mine Countermeasures missions are shown and discussed thoroughly in this article.

1. Introduction

Sonars have been used as a possible alternative to optical cameras due to the optical cameras' limitations. Sonars work under conditions which affect deeply optical cameras such as turbidity and lack of illumination. They become especially useful in underwater vehicles that lack artificial light or work too far from the bottom or in surface vehicles working in non-shallow waters, as the light attenuation in the water gives a very limited range to optical cameras.

Sonars can have several applications including but not limited to; obstacle avoidance (Karabchevsky, 2011; Petillot, Ruiz, & Lane, 2001), bathymetric mapping (Singh, Roman, Pizarro, Eustice, & Can, 2007), chain inspection (Hurtos et al., 2014b; Yong, 2011), motion estimation (Dolbec, 2007), 3*D* reconstruction of objects (Aykin & Negahdaripour, 2013) or ATR (Beaujean, Brisson, & Negahdaripour, 2011; Galceran, Djapic, Carreras, & Williams, 2012; Reed, Petillot, & Bell, 2004; Williams & Groen, 2011). Many of these applications are based on data collected with Side-Scan Sonars (SSS), Synthetic Aperture Sonar (SAS) or high resolution forward-looking sonars (FLS). Forward-looking sonars with lower resolution are also used because of their satisfactory range resolution and lower cost. Their dimensions and power requirements allow them to be mounted on Remotely Operated Vehicles (ROVs), Autonomous Underwater Vehicles (AUVs) and Autonomous Surface Vehicles (ASVs) of medium size.

One of the applications that a FLS allows is mosaicing. In the optical imaging domain, mosaicing is quite common and many examples can be found not only working in real-time (Ferreira, Veruggio, Caccia, & Bruzzone, 2012; Richmond & Rock, 2007) but also offline with higher quality both in 2*D* (Negahdaripour & Xu, 2002) and 3*D* applications (Pizarro, Eustice, & Singh, 2009). In the sonar imaging domain, there exist Commercial Off-the-Shelf (COTS) software products for post-processing and real-time mosaicing of numerous sidescan, subbottom and bathymetric sonars, such as SonarWiz5n (SonarWiz, 2013). However, in particular for FLS data, much less work has been published on mosaicing and specifically on real-time mosaicing.

Nonetheless, real-time mosaicing of FLS data can be extremely useful in applications such as Mine Countermeasures. In the context of an underwater mine detection, providing a mosaic in real-time is important to fulfill the ultimate goal of the full mission (target recognition). Typically, target recognition algorithms run on raw sonar data instead of mosaics. As it shall be seen, the Signal-to-Noise ratio (SNR) increases for mosaic data comparing it with the raw data. The mosaic can provide a better input image to the Target Detection algorithm and diminish the number of false positives, an important issue in ATR. While building a mosaic for sonar data, special care has to be taken due to the peculiarities of acoustic cameras. Nevertheless, in recent years, mosaicing algorithms for sonar data have been evolving and the current state of the art is promising.

Namely, the initial work of Kim, Neretti, and Intrator (2005) shows mosaics composed of 40 images of a boat wreckage obtained with a high-resolution FLS, Dual-Frequency Identification Sonar (DIDSON). However, the algorithm is not real-time. Later, in Kim, Neretti, and Intrator (2008), more results of a mosaic built with 80 images of a

☆ A shorter version of this paper was presented at the 19th IFAC World Congress, Cape Town, South Africa, August 24–29, 2014.

* Corresponding author.
E-mail addresses: fausto.ferreira@cmre.nato.int (F. Ferreira), vladimir.djapic@navy.mil (V. Djapic), michele.micheli@cmre.nato.int (M. Micheli), massimo.caccia@ge.issia.cnr.it (M. Caccia).

ship-hull inspection are presented. It is shown that the algorithm is implemented to work in real-time, but it is uncertain if it can provide the claimed resolution with the enhancement (up to 10 times the original) in real-time. As a result of this work, a commercial software for sonar image enhancement and mosaicing (processing time of 3.5 frames/s) is available (AcousticView, 2013). According to the authors, with this software, it is possible to obtain a mosaic of up to 1000 frames depending on the level of free memory. In comparison, our approach was tested in datasets as big as 8000 frames with no issues. Moreover, this software does not support zigzag sequences and the manual advises to perform straight lines scanning. The instructions refer to the fact that the algorithm can fail if there are no "anchor points" [sic], i.e., features to match and that a 60–70% of overlap is advisable. Another drawback is that it is specific for DIDSON, a short range and high-resolution FLS. While the range itself is not a limitation for extending the algorithm, the decrease in resolution associated to higher ranges may bring issues to the image registration. Another very recent software (SAMM, 2014), creates mosaics in real-time by stitching the images based on the GPS position without any image registration or navigation filtering. In this case, the algorithm is able to work with several different sonars. No maximum number of frames is mentioned. However, navigation filtering is only available in post-processing and no image registration is used neither in real-time nor in post-processing. In our approach, both navigation filtering and, when suitable, image registration are done online and in real-time.

Mosaicing FLS data is a very challenging task due to the approximate imaging model and commonly appearing artifacts. A very good analysis of the most important issues in mosaicing of FLS data can be found in Negahdaripour, Aykin, and Sinnarajah (2011). In Thomas, Iv, and Reed (2011), the gap produced by the nadir of the SSS is filled with FLS data. Only FLS data corresponding to the nadir of the SSS is mosaiced together with the SSS data. This method was tested in a post-tsunami survey with good results. Objects that would not be seen in the SSS data were found with the FLS. This diminishes mission time, as to see the same objects using only SSS would take more transects (due to the nadir) and thus more time. No details about computational time are given in this article. In Hurtos, Cuf', Petillot, and Salvi (2012), an innovative phase correlation-based mosaicing algorithm was applied to FLS data in a ship hull inspection scenario. The maximum number of frames registered was 834 and the algorithm took around 1 h to compute the whole set of links between the different frames. More recently, a chain inspection based on FLS small areas (4×7 m^2) mosaics was presented in Hurtos et al. (2014b). None of these were used in real-time missions. For these works, the expensive high-resolution low range DIDSON sonar is used which is not suitable for Mine Countermeasures missions where lower cost and resolution but higher range FLS are preferable. Recent evolutions of these two works can be found in the journal paper (Hurtós, Ribas, Cufí, Petillot, & Salvi, 2015). Again, the method should be able to be extended to higher ranges sonars but might suffer from the quality or lack of features. Bear in mind that, for low range and high resolution FLS like DIDSON, the size of the features found in applications such as ship-hull or chain inspection is considerably large when compared to the image size. Instead, in our work, the higher range FLS were used to image farther objects (at depths that can reach 30 m) and thus the features size is much smaller increasing the complexity. In Yong (2011), mosaicing techniques were investigated for FLS data. In that Master thesis, the algorithm works near real-time but the results are focused only on ship-hull inspection (small area covered). In this case, the maximum number of frames was 200.

As described above, the algorithms presented in the literature are not suitable to work for a wide range of applications, in real-time and for large scale areas. The state of the art methods try to solve a specific problem and are not focused on the real-time constraint. The work presented here overcomes all these limitations. It tries to be as generic as possible while maintaining the real-time constraint and working in any area of any dimension. It generates georeferenced mosaics that can be easily overlapped in a satellite map. Namely, it can be considered for ATR applications, large area survey and post-mission analysis, among others. The algorithm is flexible to work with various sonars (BlueView and Reson tested thus far) and in different setups (fixed to a pier or mounted onto a moving ASV or onto a moving ROV tested so far). The results presented here are more related with the application of Mine Countermeasures, specifically ATR. For more results on large scale areas please consult (Ferreira, Djapic, & Caccia, 2015). Due to the few works found in the literature regarding sonar mosaicing, the reasons that motivate this line of research are introduced in the next section together with its application to Mine Countermeasures. A description of the mosaicing algorithm follows in Section 3. Section 4 describes the target recognition algorithm. The results are presented and discussed in Section 5. Finally, Section 6 concludes the article and proposes future work.

2. Motivation and applications

2.1. Motivation

Forward looking sonars have seen an impressive technological development in the past few years with higher frequency sonars in the range of MHz. Some of the latest commercial forward looking sonars can go over 1 MHz up to 3 MHz providing high quality images for ranges between few metres and dozens of metres. This allows new applications such as sonar-aided navigation (Johannsson, Kaess, Englot, Hover, & Leonard, 2010), chain inspection (Hurtos et al., 2014b) or mosaicing. Due to their high quality, sometimes, forward-looking sonars are named acoustic cameras. In what follows, these two terms are interchangeable.

There are several reasons that motivate the mosaicing of FLS data. One of them is the filtering of the acoustic noise. Reducing the noise increases the SNR. This happens because of the averaging effect involved in the mosaicing process. Comparing with optical cameras, acoustic cameras have intrinsically more noise due to the physics of the image formation. For optical systems, the experimental conditions can be defined in a way that minimizes noise (e.g., using homogeneous illumination). In the acoustic domain, the noise is considerable and mosaics can reduce it significantly. Defining favourable experiments like mounting the sonar in an overactuated stable robot can only help to reduce the influence of perturbations. However, this is not enough to diminish the noise sufficiently. The physics of an acoustic device such as a forward looking sonar implies that several consecutive images will not be very similar. For instance, backscatter and reflections coming from the water column occur independently of the stability of the platform and affect the data quality.

For surface vehicles, the sea state influences the sonar noise level. Waves can have a considerable impact in pitch and roll variations. These degrees of freedom are not controlled normally. Their instability affects the insonified area and incident angle. Fig. 1 exemplifies this inhomogeneous insonification natural phenomena. It shows two almost consecutive frames (separated by one frame and half a second) with different insonifications even though the vehicle is practically in the same place. This can severely affect the performance of object detection algorithms as an object can be seen clearly or hardly depending on the insonification. Several frames with different insonifications are normalised by the averaging process implied in the building of a mosaic, solving that issue.

Other acoustic devices constitute also sources of acoustic noise. Namely, echosounders and DVL can interfere with a FLS when mounted onto the same vehicle and working in frequencies within the operational range of the FLS. The same averaging effect of mosaicing can substantially decrease this source of noise. To better understand this issue, real data collected during sea trials exemplifies

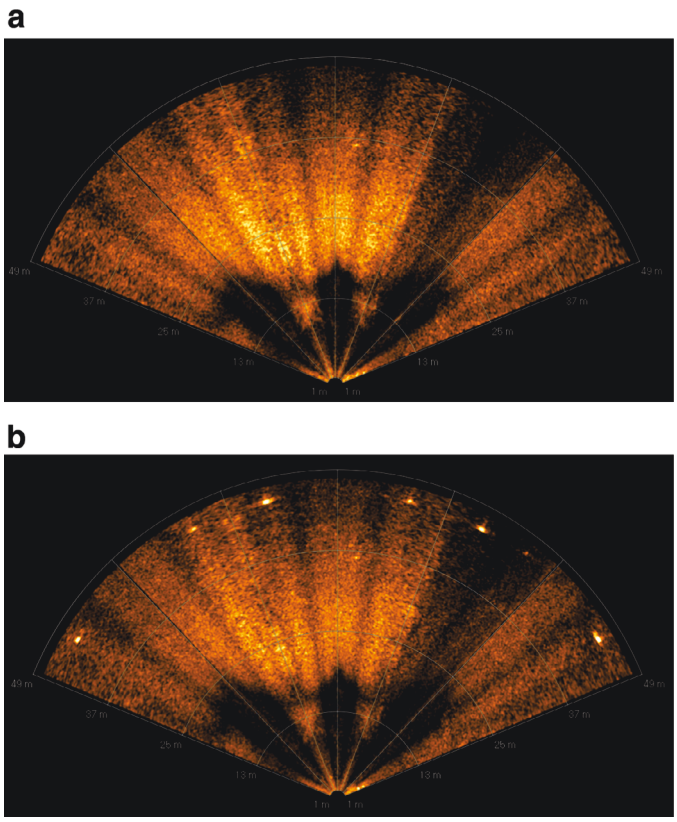


Fig. 1. Two frames obtained at close times (0.5 s difference) with different insonifications at a range of 50 m.

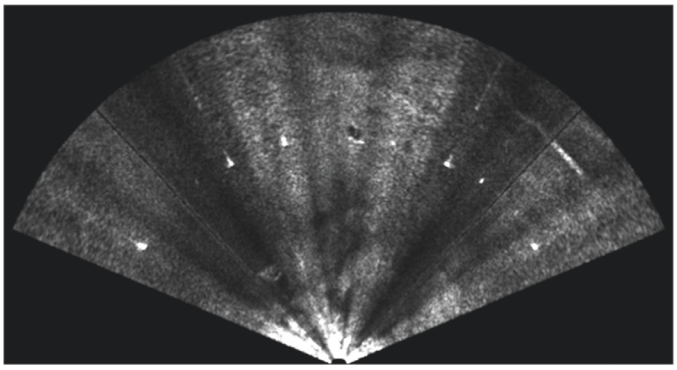


Fig. 2. Raw data showing a target in the middle and echosounder interferences to the left and right of the target.

this situation in Fig. 2. This is an example of the interference that an echosounder can introduce. For this particular case, the echosounder works at 200 kHz while the operating frequency of the FLS is 450 kHz and the range is set to 50 m. The BlueView P450-130 sonar also detects echos between 300 kHz and 600 kHz. In Fig. 2, the reader can see very bright interferences coming from the second harmonic of the echosounder (400 kHz). This sonar is made of six heads and thus six bright blobs can be noticed, one per each sonar head. As it will be seen in the Results section, this interference provokes a considerable number of false positives in an Automatic Target Recognition (ATR) application. This happens because the echo coming from the echosounder harmonic is similar and brighter than the one coming from the target to be detected and recognised.

Another acoustic device that can interfere is the DVL. In our experiments, it revealed to be less problematic than the echosounder as its operating frequency was 600 kHz. This is the limit frequency that the

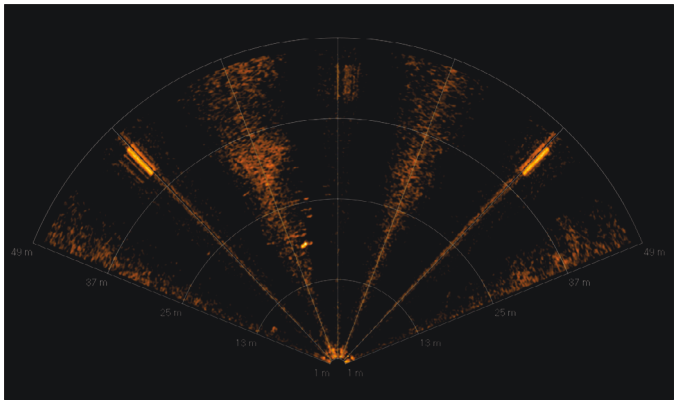


Fig. 3. An example of the interference provoked by the DVL. Note the 3 rectangular shapes in the range direction on the top left, middle and right. Range of the sonar is 50 m.

BlueView P450-130 can detect. Fig. 3 shows the kind of interference the DVL can provoke. It is less problematic when compared to the echosounder mainly because of the way the sonar is built, its operating frequency and kind of signal. Without going into details, only 3 DVL-related blobs are present for a given image as it can be seen in Fig. 3. Moreover, due to its bigger pulse length, the DVL creates longer patterns in the range direction which, in turn, makes the shape of the interference completely different of the objects to be recognised. However, if bigger/other shape objects were considered, the DVL interference would be more harmful. The algorithm presented here aims to be as generic as possible. Therefore, regardless of the application, eliminating the DVL interference is another reason to mosaic the data.

Finally, a very important reason to mosaic forward looking sonar data is the improvement at the image contrast and detail level. Image contrast improves clearly with the averaging effect of the mosaicing process. With a better contrast, both man-made and natural objects present in the seabed are much more easily identified. In the raw data, an object can be occluded and not visible for some of the frames. Moreover, it is hard to identify details that cross several frames. Instead, in a mosaic, even if the object is not seen for a particular frame, it can be easily identified in the next one. Due to the stitching of several frames, an object is insonified more than once and it can be seen in multiple views. The final result is that more details can be observed in the mosaic and long continuous objects are well identified. For instance, long mooring cables can be hard to recognise in raw data with partial views while they are very clear on a mosaic. The more detailed mosaic has also a better resulting contrast than each single image.

Mosaicing forward looking sonar is useful for the reasons enunciated above. Its performance in real-time can be extremely helpful to evaluate a mission and further plan future missions or possibly re-plan the missions in real-time. Human operator can immediately inspect the surveyed area while the robot is moving without waiting until the mission's end. Moreover, an operator can define previously when to inspect partial mosaics by defining start and stop geographic coordinates that correspond to a transect in a lawn mowing pattern for instance. In particular for FLS data, this can be important to avoid artifacts.

2.2. Application to Mine Countermeasures

There are several applications for the mosaicing of FLS data including chain inspection, survey of large areas and object detection. The reader can refer to Ferreira et al. (2015) for large scale areas mosaicing and its motivation. Hereby, the specific application of Mine Countermeasures in general, and Automatic Target Recognition (ATR) in particular, are introduced.

As it was described in the previous subsection, mosaics improve contrast with respect to raw data. This enhancement is reflected in the performance of an ATR system running on mosaiced data instead of raw data. In a typical mission, a list of possible targets is given by an AUV equipped with a high-resolution SAS or SSS system to an ASV with a lower resolution FLS. The list of targets can include information such as the targets' estimated position, their orientation and their shape. The ASV performs a closer inspection in the area around the expected target location. The inspection can be performed with different motion patterns, namely circling around the target, keeping the distance to it or a cross-pattern that makes perpendicular and parallel tracks to the target. For any of these patterns, partial mosaics can be built and used as input of the target recognition algorithm. Target recognition algorithms run typically on raw data. However, by using mosaics as input, the false positive rate diminishes as it will be shown in the Results section. Special care has to be taken when considering mosaics for ATR, e.g., in circular missions where the shape of the targets can change considerably. For cylindrical-shaped targets, both the shape and the shadow vary with the angle in such a way that a target recognition algorithm has to deal with the several different cases. Thus, instead of a full mosaic for a 360° turn, partial mosaics that cover an area where the shape does not change greatly have to be produced. This ensures a consistent representation of the target. These mosaics are then provided as input to a target recognition algorithm. The mosaicing algorithm presented here is connected to the ATR algorithm developed in Galceran et al. (2012) and recently updated in Ferreira, Djapic, Micheli, and Caccia (2014).

In the context of ATR, the Post-Mission Analysis (PMA) time can decrease greatly if good detections are provided by running ATR on mosaics. Current state of the art AUVs used in Mine Countermeasures missions are pre-programmed. The robot collects the data using its onboard sensors (side-scan, and recently forward-looking sonars used as gap fillers) and the operators involvement is required for PMA of the sonar data. Typically, human operators spend hours analysing the waterfall display of the data in order to find potential mine-like objects. This operation tends to be very exhausting. The operators are prone to make unintentional mistakes after some hours of such an operation. Nonetheless, nowadays, human operators still outperform software ATR algorithms. Even if the probability of detection can be close to the human operators, computer ATR algorithms tend to have much greater false-alarm rates. This is undesirable if some additional search patterns are implemented to take a closer look at the computer ATR targets as they can dramatically increase the overall mission time. Namely, if the number of false positives on the initial search is much greater than the real targets, any autonomy search pattern, even if maximally optimized (travelling salesman, etc.), will further increase the already long mission time as the ATR results imply that the AUV must take many "second looks." Thus, the decrease in false-alarm rates by using mosaics instead of raw data is very important in this context. Diminishing the false-alarm rates avoids extra inspection missions which decreases the overall mission time.

Finally, still in the context of ATR, mosaics can be useful to distinguish between a bottom target and a moving target. For instance, for the Mine Countermeasures missions considered here, after the recognition of a target by the FLS onboard the ASV, a small Unmanned Underwater Vehicle (UUV) is launched by the ASV. This UUV has to reach the bottom target previously detected by the ATR algorithm running on the mosaics. The UUV is detected after the launch and then tracked along its way until it reaches the bottom target by analysing the FLS raw data. The guiding of the vehicle until the bottom target is explained in Miskovic, Djapic, Nad, and Vukic (2011) and it is based on navigation updates sent acoustically by the ASV to the UUV. One of the issues that can occur in this mission is a situation where the automatic tracker of the UUV fails and the UUV gets lost as the received navigation updates are wrong. This can happen commonly in two situations: 1) either the tracker gets another blob in the image such as sea grass (Posidonia) or rocks when the UUV is passing by one, or, 2) the tracker jumps to the bottom target when the UUV is reaching it. In both cases, the mosaic images can be very useful not only in post-processing but also in real-time. Namely, the latter case is very hard to identify in the raw images in real-time. When the UUV is very close to the bottom target, the vehicle and the target look like one single blob. If a mosaic of this mission phase is built, then it is very easy to see in real-time in the mosaic that the robot has passed the bottom target. A mosaic can be useful especially if background subtraction is implemented, as it shall be described in Section 5.4.

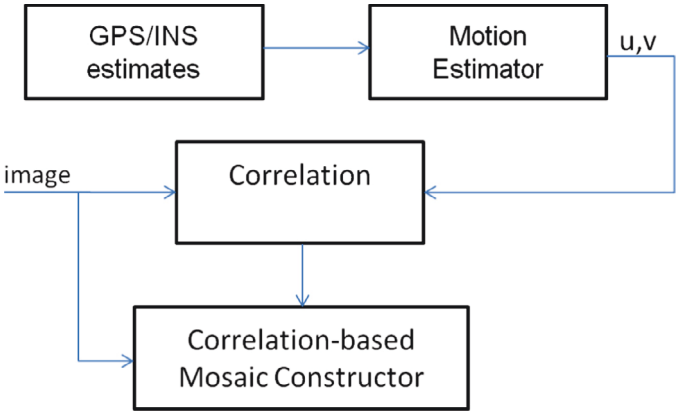


Fig. 4. Correlation-based sonar mosaicing architecture.

3. Description

The general architecture for the forward looking sonar mosaicing (acoustic camera) is inspired in Ferreira et al. (2012) and shown in Fig. 4. That work describes a vision-based real-time mosaicing algorithm that used the Simultaneous Localisation and Mapping (SLAM) output for optical image registration purposes. However, due to the use of acoustic images, its specific application of Mine Countermeasures and the tested datasets, SLAM is not included as a basis for the mosaicing process here. Instead, correlation is used to correct the motion estimation guess and complete the registration process. This replacement is due to the lack of good features for the loop closing. Closing the loop with false positives can produce substantially wrong mosaics. It is better to have some drift but make sure the data is consistent than risking getting very inconsistent results because of the false positives in the data association SLAM process. This is object of future work as the final goal is to include SLAM in the mosaicing process.

The other major difference is that with acoustic cameras, it might be very hard to estimate the motion of the vehicle for the same reason. The lack of good features to track makes sonar odometry a very difficult task. While for optical cameras, a vision-based motion estimation was performed, here, the motion can be estimated recurring to the navigation sensors of the vehicle. Namely, for an ASV, GPS or DGPS (in our case) estimates are available for the motion estimation. For a ROV, different options exist for its navigation and localisation including Long BaseLine (LBL), Ultra-Short BaseLine (USBL), DVL, Inertial Navigational Systems (INS), Inertial Measurement Units (IMUs), etc. In our particular case, the ROV is equipped with a PHINS INS that integrates a high-precision IMU, DVL and USBL measurements. Similarly, instead of using a magnetic compass, heading information is available from the DGPS system or the PHINS unit.

In order to relate the motion estimated by the robot to sonar coordinates, an image formation model is needed. In a forward looking sonar insonifying the sea bottom, the 3D points of the insonified area can be represented in spherical coordinates (r, θ, ϕ). Fig. 5 shows the sonar image geometry. For a matter of simplicity and without loss of

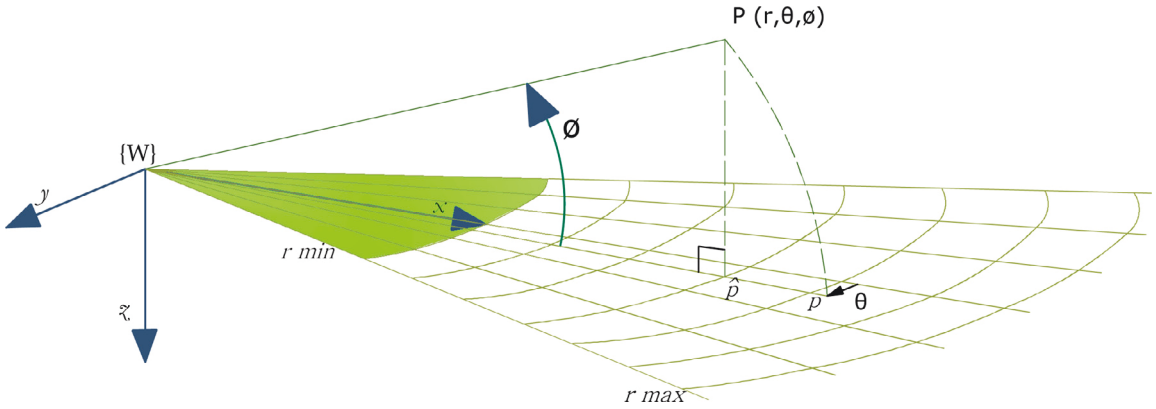


Fig. 5. Imaging sonar projection model.

generality, the sonar centre is the world reference frame. The range of the sonar r, the bearing angle θ and the elevation angle ϕ define the 3D coordinates of a cartesian point $\mathbf{P}$ as shown in Eq. (2). The conversion between cartesian coordinates to spherical coordinates is as in Eq. (1).

$$\mathbf{P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r\cos\theta\cos\phi \\ r\sin\theta\cos\phi \\ r\sin\phi \end{bmatrix} \tag{1}$$

$$\mathbf{S} = \begin{bmatrix} r \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{x^2+y^2+z^2} \\ \tan^{-1}(x/y) \\ \tan^{-1}(z/(\sqrt{x^2+y^2})) \end{bmatrix} \tag{2}$$

This model is not very useful to process the data. The raw data that the sonar provides includes only slant range and bearing. Therefore, an approximative model is needed that translates 3D points into 2D points. A forward looking sonar is normally mounted with a certain tilt angle illuminating the seafloor. Generally, this tilt angle is small (5°). For typical FLS, the vertical beam width of the sonar is about 10°. These two conditions are necessary to make a 2D approximation of the real 3D scene. As in similar work in the literature (Hurtos et al., 2012), the 3D point $\mathbf{P}$ is projected onto the 2D point $\mathbf{p}$. Eq. (3) presents the projected 2D point $\hat{\mathbf{p}}$ using an orthographic projection as approximation with m and n the approximated x and y coordinates. This approximation is valid as long as the scene relief is negligible when compared to the sonar range. This applies if the two conditions above are true. In that case, the elevation angle ϕ is small making $\cos(\phi) \approx 1$ and $\sin(\phi) \approx 0$ validating Eq. (3).

$$\hat{p} = \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} \tag{3}$$

Considering the orthographic projection, an affine transformation can be used to relate two consecutive frames. The rotation matrix and translation parameters for the affine transformation can be directly obtained from the navigation systems of the vehicles that carry the FLS. In the case of the ASV, the DGPS system can provide the translation vector and the heading variation. There is a coordinate transformation from the DGPS receiver position to the sonar head. Additionally, the sonar head is mounted on a variable depth pole and has a pan-and-tilt unit. The depth of the pole only matters when georeferencing the mosaic as it induces a difference in the slant range with respect to the surface. Instead, the pan-and-tilt unit has to be considered in the affine transformation. As written above, as long as the tilt angle is small, the approximation holds. The pan angle is treated as a pure rotation around the elevation axis. In case of the ROV, the sonar was fixed but the robot was fully actuated. It was able to control the six translational and rotational degrees of freedom (DoF). The pitch angle was purposely kept at 0° and the robot depth was controlled to obtain optimal sonar position with respect to the sea bottom. The robot was also stable in roll. The ROV integrated navigational system provides the translation vector and rotation angle. Eq. (4) represents the affine transformation with (m, n) being the pixel coordinates of a projected point. For a matter of convenience, Eq. (4) can be rewritten in homogeneous coordinates for the projected point (m, n, 1). The final formula is shown in Eq. (5) with $\mathbf{R}$, rotation matrix and $\mathbf{t}$, translation vector.

$$\begin{bmatrix} m_2 \\ n_2 \end{bmatrix} = \mathbf{R}\begin{bmatrix} m_1 \\ n_1 \end{bmatrix} + \mathbf{t} \tag{4}$$

$$\begin{bmatrix} m_2 \\ n_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0_{1x2} & 1 \end{bmatrix}\begin{bmatrix} m_1 \\ n_1 \\ 1 \end{bmatrix} \tag{5}$$

The effect of this approximation in the image registration process was studied in Johannsson et al. (2010). Commonly, this error is in the order of centimeters. Other possible sources of errors are the changes in roll but they can be considered negligible. On the other hand, changes in pitch are normally related to the tilt angle. As mentioned above, for small tilt angles the approximation is valid. The change of the tilt angle only changes the reflected intensities and the border of the insonified area. Thus, respecting the assumptions referred above, one can use an affine transformation to relate two sonar images.

As mentioned above, the motion estimate cannot be used in a SLAM framework in featureless environments. Hence, the motion estimate can be used directly as the initial guess of the point where the mosaic and the current image should be stitched together, i.e. the point given by the affine transformation. Then, the same correlation method as in Ferreira et al. (2012) can be applied in a neighbourhood of that initial guess. The method uses the normalized correlation coefficient computed in a neighbourhood of 20 × 20 pixels. In some cases, e.g., small featureless areas such as the ones present in some Mine Countermeasures missions, the need for this refining step drops. In these missions, the robot moves slowly when compared to the frame rate obtained and thus the DGPS/ROV navigation estimation can suffice. Moreover, normally, one of the few features present in such missions is the target to be detected and this is not enough to use correlation for the whole mosaic. In this kind of missions with very few features, the mosaics produced are relatively small. Thus, they do not suffer considerably from drift and using filtered good navigation data can be an option.

In featureless environments, the correlation method can have poor performance. For instance, in sandy bottom areas with no artificial or natural objects of interest, using correlation to refine the location of the best point to stitch together mosaic and current frame is not always beneficial. If one applies a feature-based method such as correlation, the algorithm estimates that the robot is practically stopped as the actual image is very similar to the mosaic leading to

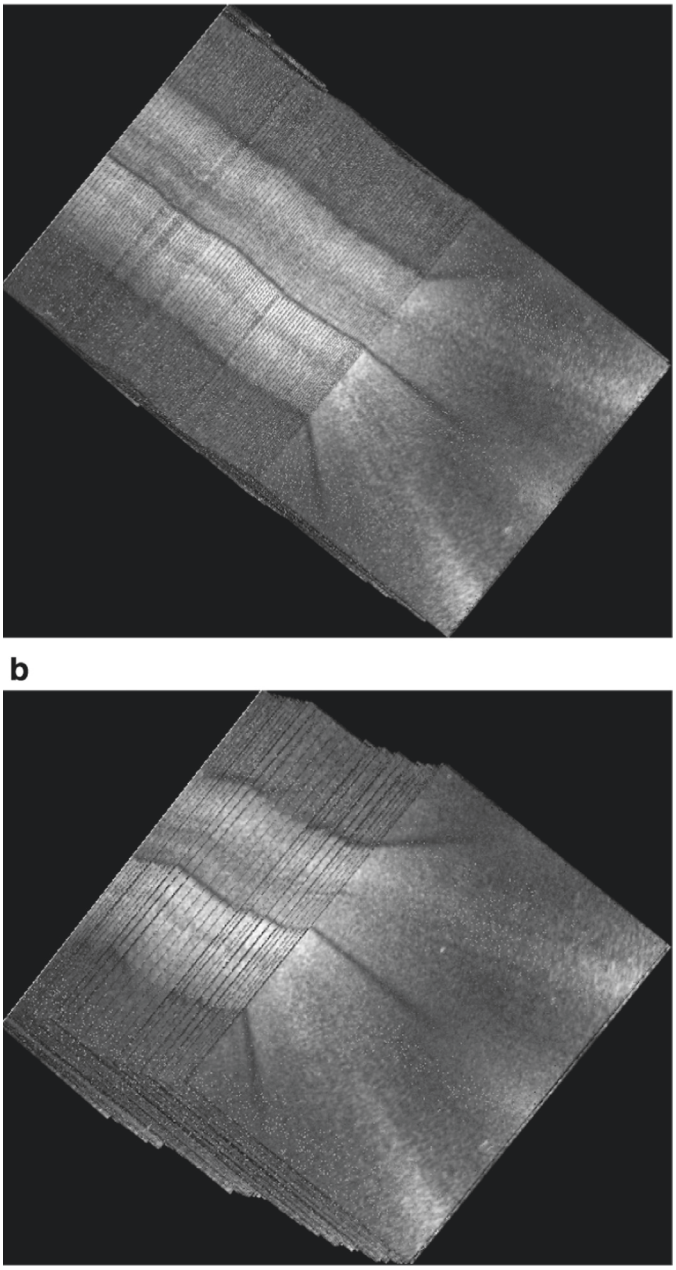


Fig. 6. On the top, a partial of the mosaic built only with DGPS estimates. On the bottom, a partial of the mosaic of the same area using also the correlation method.

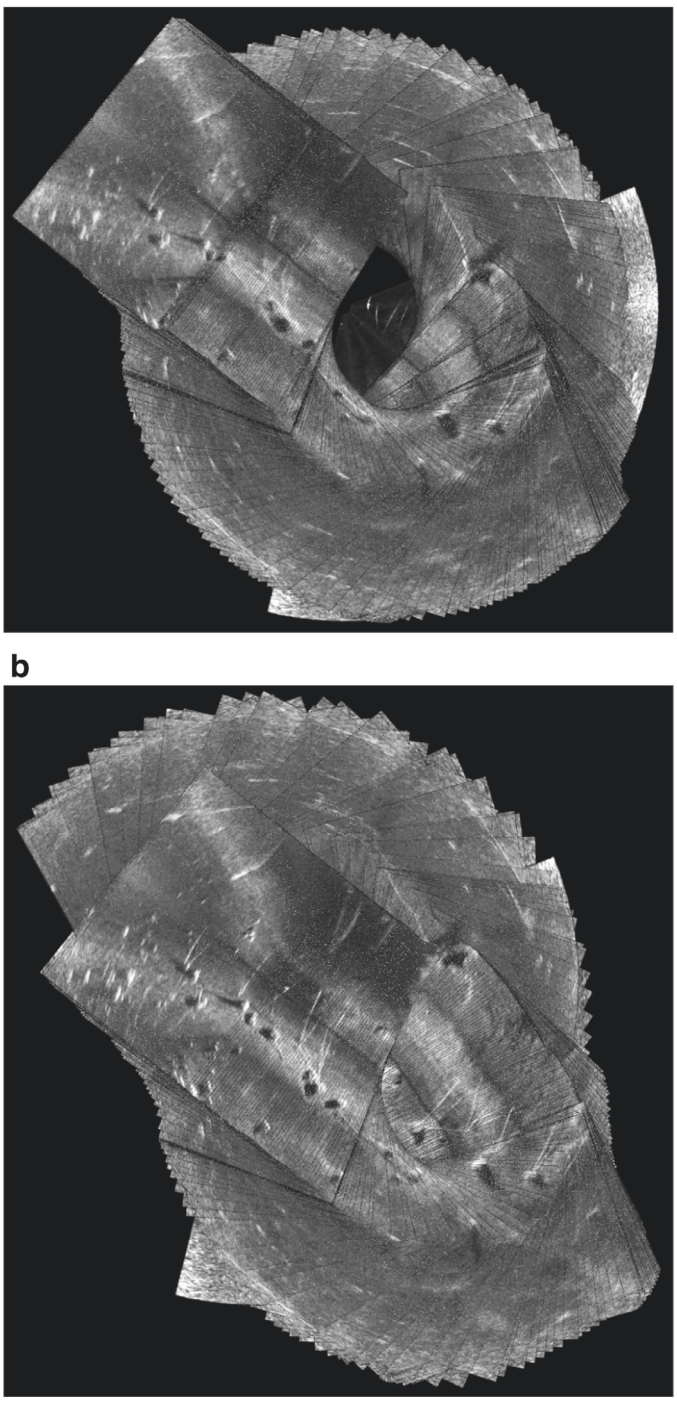


Fig. 7. On the top, a mosaic built only with DGPS estimates. On the bottom, the same mosaic using the correlation method. The number of frames used is around 700.

a maximum correlation coefficient for a zero pixel translation. Fig. 6 shows an example of this situation with the same area represented in a mosaic that only uses DGPS estimates and one that uses the correlation method. As it can be easily seen on Fig. 6b, several frames map to the same location if the correlation technique is applied. There is also an error in the direction perpendicular to the movement of the robot due to the noise sensitivity of the correlation method to very similar images. For the BlueView sonar tested, two sonar heads are joined together and a thin darker line departing from the origin represents the border of each sonar head. These lines should not considerably affect the correlation method as they fade towards the top of the frame and the correlation is computed in area far enough from the sonar head. The lack of continuity of these lines means that the algorithm is drifting as the robot is actually only moving forward and thus the mosaic should show a straight line. In Fig. 6a where only DGPS estimates are available, these lines are continuous. In these situations, using only DGPS gives a better result than the correlation method. However, filtering of the DGPS estimates is needed to avoid possible jumps on the GPS position. Both maximum speed and acceleration of the vehicle are taken into account in this filter. Current work tries to identify environments that are featureless, and switch between using correlation or only DGPS (or other on-board navigation sensors for underwater vehicles).

In environments with a sufficient number of features, the correlation method improves the quality of the mosaic and corrects for the navigation drift. Fig. 7 is a good example of the effectiveness of

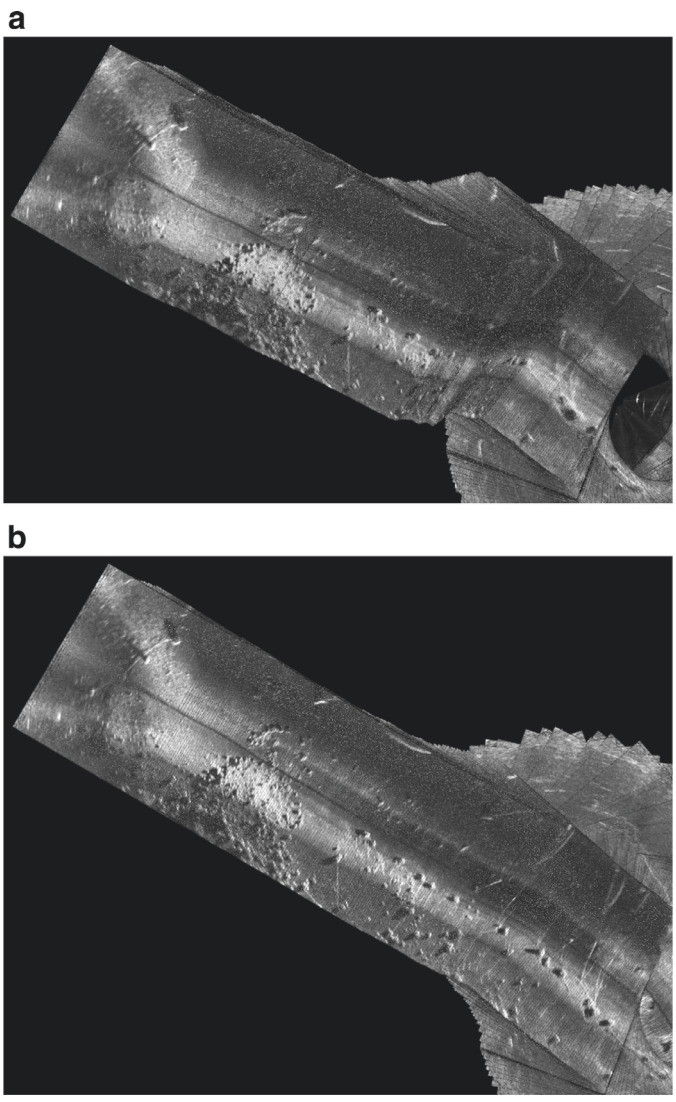


Fig. 8. On the top, a mosaic built only with DGPS estimates. On the bottom, the same mosaic using the correlation method. The number of frames used is around 1070.

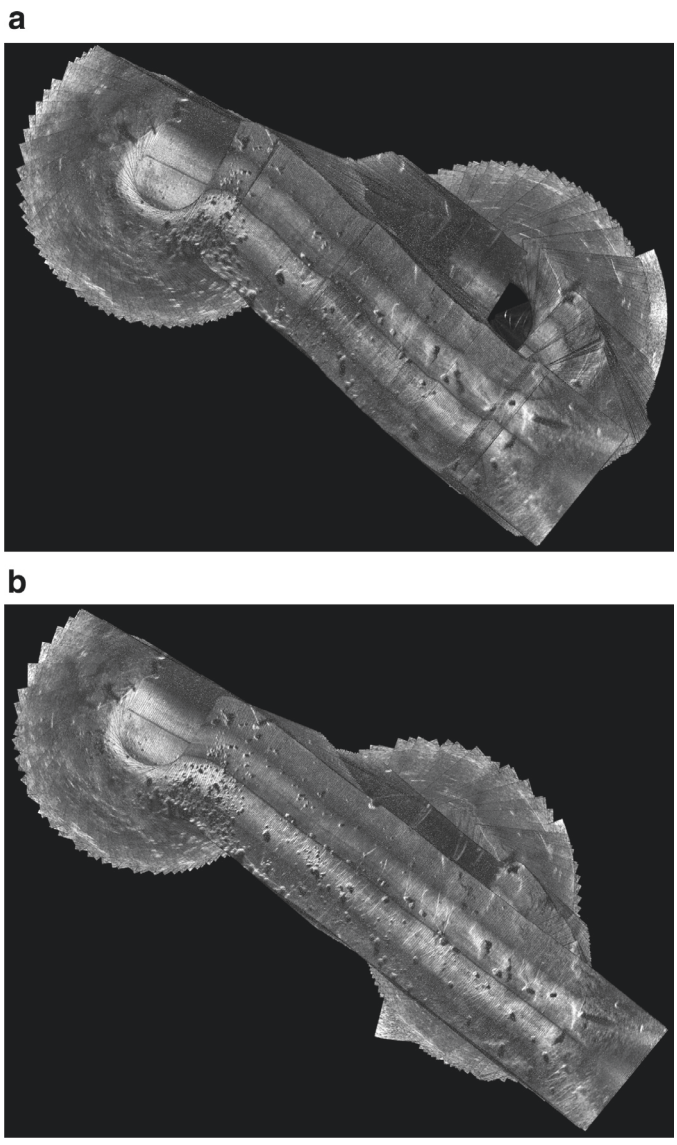


Fig. 9. On the top, a mosaic built only with DGPS estimates. On the bottom, the same mosaic using the correlation method. The number of frames used is around 1700.

the correlation method in the correction of the navigation drift. This figure shows an example of a mosaic built using around 700 frames both without the correlation method and with it. It is very clear that Fig. 7a has several frames mapping to the same place (top left) due to issues with the navigation. Instead, in Fig. 7b, the correlation method corrects the navigational errors by registering correctly the images.

To understand better the improvement, a bigger mosaic with 1070 frames is shown in Fig. 8. This mosaic includes the one from the previous figure so it can be seen the effect of previous navigational errors in the part of the mosaic that is created afterwards. In Fig. 8a, besides the same problem of Fig. 7a, the drift is clear and has the effect of diminishing the length of the transect. With respect to Fig. 8b, which includes registration by the correlation method, the erroneous mosaic is around 22 m short in a 200 m transect, thus over 10% of error.

Finally, a mosaic from the same sequence built from 1700 frames is presented in Fig. 9. Due to several accumulated drifts, e.g. the one from Fig. 8a can be identified in the centre of Fig. 9a or the mapping to the same place in the bottom right, Fig. 9a shows a poor result of a mosaic built with navigation data only. Instead, Fig. 9b represents the same area in a seamless mosaic that uses correlation to register the frames.

Mosaic building and blending: A simple image registration that refines the DGPS (or INS) estimates using correlation updates the mosaic. For the building of the mosaic instead, special care has to be taken. Acoustic data is not as linear as optical data and the intrinsic subtleties derived from the image formation model seed the need of a particular blending.

First of all, in a forward looking sonar, the raw image produced by the sonar does not represent a full 2*D* image of the sea bottom as it happens in a downward-looking optical camera. Indeed, due to the positioning of the sonar, the raw data captured by the acoustic camera also includes reflections from the water column. Most of the pixels close to the origin of the sonar head should be discarded. This is because they do not represent accurately the bottom. Instead, the closest pixels to the sonar head include water column reflections. Moreover, the borders to the left and right are also cut as normally the survey is performed in a way that the area of interest is in front of the sonar head. Therefore, only a valid region of interest should be mosaiced. This should be defined taking into account the expected depth of the area and the tilt of the sonar so that useful pixels (i.e. representing the bottom) are not cut off. Exceptionally, when the goal is to mosaic static setups or setups where a moving target is actually close to

the sonar head as presented in Djapic et al. (2013), then an extended area, covering most of the sonar's Field Of View (FOV) should be considered. Another issue that can be problematic in certain application scenarios such as Mine Countermeasures is the shadows provoked by objects with a certain height. The size of the shadow depends mainly on the height of the object, its shape, range to the sonar and tilt angle. This shadow also changes with the azimuth angle of incidence as well as the shape of the object. For applications related to target detection and recognition, this can be a problem as the object and shadow look differently depending on the view point. What can be done in those cases is to build partial mosaics instead of a full mosaic of the whole mission. For instance, in a cross-pattern mission where a parallel and perpendicular transects (with respect to the object) are used, separate mosaics for parallel and perpendicular transects should be built. In a circular mission where a robot circles around the object to be recognised, this issue can be even more notorious if the shape of the object changes considerably with the viewpoint (e.g. a cylinder). Thus, in such missions, the algorithm builds partial mosaics covering only portions of the full 360° where the object's shape and shadow do not change drastically.

The blending strategy cannot be the same as the one used for optical cameras. In that case, only new pixels were added to regions that were empty. For FLS data that approach does not apply. First, in static setups where there is a moving target, no update would occur and the history of the mission would be lost. Second, in circular missions around an object, that would mean adding only pixels in the boundaries and not filtering any noise in the overlapped areas. For lawn mowing patterns that strategy can be used. However, it is a suboptimal strategy, as it does not improve areas that are seen in more than one frame whereas it is exactly for those cases that the mosaic can be useful. Due to changes in insonification, an object can be clearly seen in one frame and then not anymore in the next one(s). If the blending technique would add only new pixels in empty regions and the object was not seen on the previous frame on the imaged region, then that object could not be seen at all in the mosaic. Hence, a different approach is hereby used. Instead of a superimposing method, the blending is based on an average of both images. The overlapping area between the mosaic and the current frame is computed. For this area, an average of the mosaic and the new frame is copied to the mosaic. This has a noise filtering effect and gives more importance to the new frame than what it had in the optical camera mosaicing. This simplified weighting technique works sufficiently well for many applications. Other more complex techniques exist. For instance, the ones described in Yong (2011) or recently in Hurtos, Cufi, and Salvi (2013) include different weights for each image or for each image region. These techniques can be used in a nonreal-time context but, in this case, the real-time constraint led to the choice of a simpler yet effective solution.

4. Target recognition algorithm

The mosaicing algorithm output is connected to the target detection and recognition algorithm described in Galceran et al. (2012). This algorithm was updated in Ferreira et al. (2014) with several improvements. Originally, it is a real-time detection algorithm for FLS raw data that makes use of integral image representation to achieve the real-time capability. Targets are detected by comparing the echo map of a region of interest and the background map of the same area. Only pixels that have an echo a certain amount higher than the background (threshold configured) are considered as possible targets. The resulting blobs after thresholding are morphologically analysed taking advantage of prior knowledge about the kind of objects that are being looked for. Then, a minimum echo threshold is used to filter out lower intensity blobs. The target expected location is given to the target detection algorithm by the results of a survey using SAS. In the case of a cylinder, its orientation can be estimated from the SAS data and provided as an input to the detection algorithm. Only location and orientation are provided by the SAS survey not the full map.

Having the orientation of the target, one can search for the cylinder only when its broadside is visible. This is especially important in the case of circular missions around the target as its shape changes with the relative bearing to the sonar. The maximum deviation from the ideal case where the cylinder is perfectly horizontal on the sonar frame is parameter configurable and it ranged from 30° to 50° in the experiments. The final decision of choosing the target location from the several detections produced was also improved. The detections are grouped and the distance from the centroid of each group to the expected target location is used. In this way, the closer a cluster of detections is to the target, the most probable is to be chosen. The Euclidean distance between the shape of each detection and the ideal shape is calculated and averaged. Then, the score obtained for each cluster is weighted with the inverse of the mean distance. In this way, a cluster that has less detections but whose detections are closer to the real object increases its chances of being chosen.

Due to the modularity of the MOOS framework (Newman, 2009) framework, very few changes were needed to connect the output of the mosaicing algorithm with this algorithm. Both the mosaicing and ATR run on a MOOS-IvP (Interval Programming) (Benjamin, Schmidt, Newman, & Leonard, 2010) environment that controls in real-time the full mission of a multi-robot system. The target detection can be enabled by the mosaicing algorithm when this produces and publishes a new mosaic. After performing a detection, the target detection algorithm will wait for a new mosaic to analyse. The detections are all saved in a list and the final decision is taken after either a certain amount of time, number of mosaics analysed or by request. How often the mosaicing algorithm publishes a mosaic is configurable and has to do with the type of mission.

5. Results

The results presented here regard mainly the application of mosaicing to Mine Countermeasures. They are not exhaustive of all the possible applications for this algorithm. For more details on the mosaicing of large scale areas, please refer to Ferreira et al. (2015). In Ferreira et al. (2015), a mosaic built from a 42 min mission covering an area of 75,000 m^2 and over 8000 frames is presented. The length of each transect was in the order of 500 m. Here instead, smaller mosaics are presented as for the application of Mine Countermeasures Reacquisition, the typical missions are focused on smaller areas. Nonetheless, one of the mosaics presented in this section was built from 2511 frames in an area of about 3000 m^2. All the results were obtained either by processing previously collected data or by working with live sonars during real missions. No simulated results are presented.

The data was collected by different vehicles equipped with various sonars working in different conditions during several experimental trials. Namely, a BlueView sonar model P900-130 working at 900 kHz and with 130° FOV mounted onboard an ASV, a BlueView P450-45 (45° FOV) fixed to a pier and a Reson SeaBat 7128 working at 400 Hz, 128° FOV onboard a ROV. The algorithm can run at a frame rate between 1 Hz and 4 Hz depending mainly on the number of mosaics saved for post-processing and data analysis. The neighbourhood for correlation computation is small and thus the algorithm can register the image very fast. As it shall be seen, in some cases, a considerable amount of data can be useful for post-processing and thus it is saved during the real-time operation which can decrease the frame rate. However, a minimum of 1 Hz was possible to maintain at all times. In the next subsections, details about each trial, application and sonar used will be given. Section 5.1 gives examples of the improvement mosaicing brings to object contrast. Then, Section 5.2 presents a full mosaic of an ATR mission. The Automatic Target Recognition application itself is detailed in Section 5.3. Section 5.4 discusses how the mo-

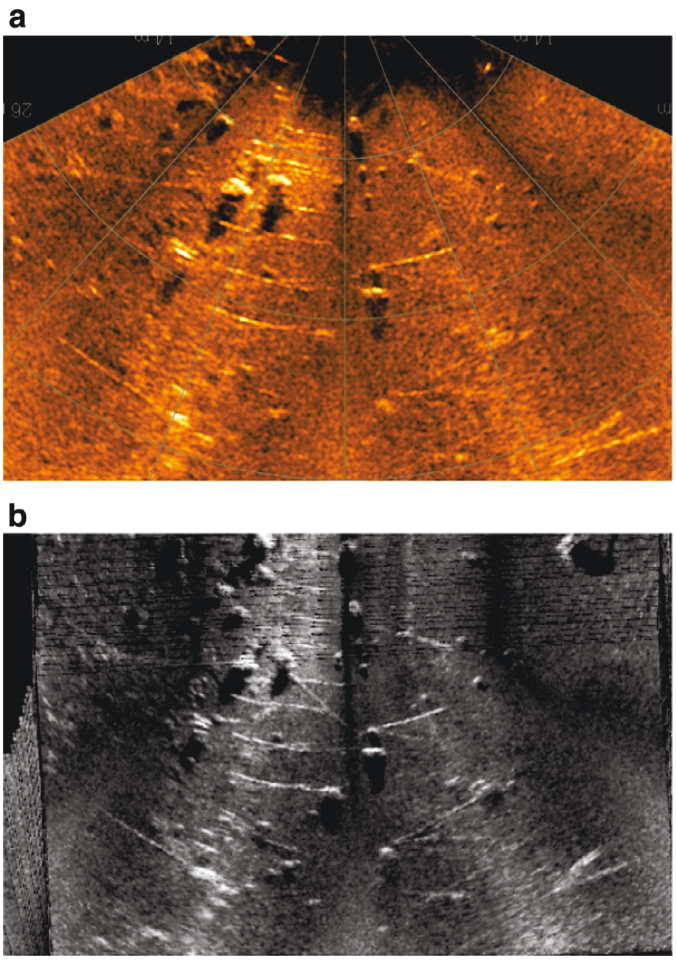


Fig. 10. On the top, the raw data. On the bottom, the mosaic for the same area after 1412 frames.

saics can be useful for detecting moving targets either in real-time or post-processing. The fact that the mosaics are built in real-time does not prevent their use in post-processing as it shall be seen in this section. Finally, Section 5.5 shows the importance of the data size reduction provided by mosaicing in the context of Mine Countermeasures missions.

5.1. Object contrast improvement

This section concentrates on the improvement in object contrast that mosaics bring when compared to raw data. This improvement is the main motivation to the use of mosaics for the purposes of ATR and thus is here introduced. The data presented in this section was obtained with a BlueView P900-130 sonar mounted onboard the ASV Gemellina (modified 4 m Sea Robotics USV). The range of the sonar was typically maintained between 30 m and 50 m for all the experiments with the BlueView sonars in this section and the following ones. The range resolution is around 5 cm. The data was collected in Marciana Marina, a marina in the Island of Elba, Italy during the ANT11 trial conducted by the NATO STO CMRE, formerly known as NATO Undersea Research Centre (NURC). Rather than the full mosaic, detailed views of portions of the area covered can introduce better the improvement and usefulness given by mosaicing. Fig. 10 shows partial views of both raw data and the mosaic built until that point. The shown range is around 30 m for the raw data and slightly more for the mosaic. As it can be easily observed, the mosaic provides a better contrast between an object and the background. Not only that, it defines better the object's shadow. Finally, it allows the user to see

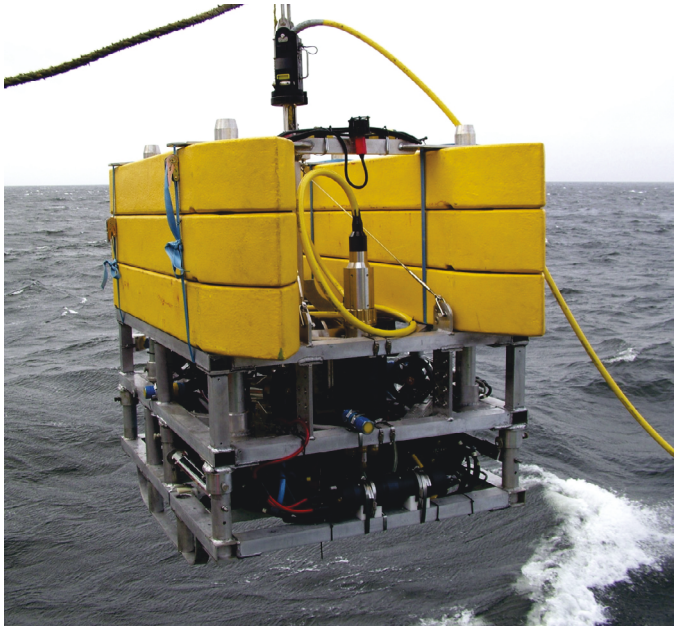
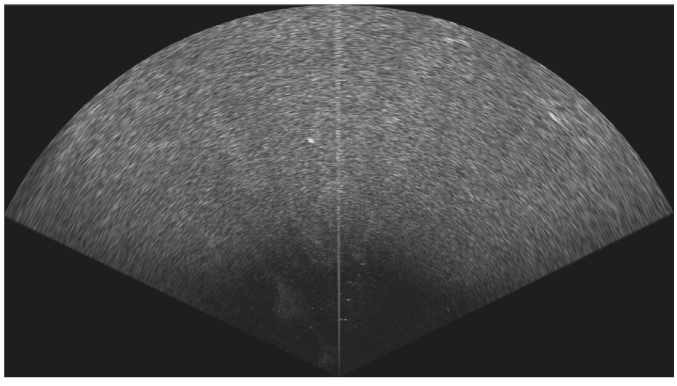
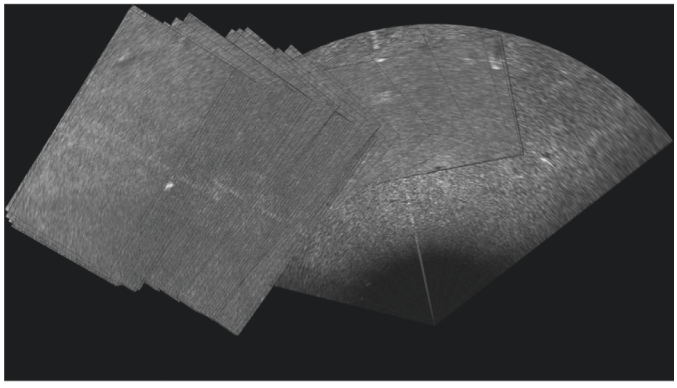


Fig. 11. The ROV Latis from University of Limerick.



(a) The raw data.



(b) The mosaic.

Fig. 12. The same target observed in a mosaic and in the raw data. The mosaic is built from 73 frames.

details hardly seen on the raw data such as the mooring ropes. In the raw data, it is hard to picture the full mooring ropes present in the scene. At the same time, the SNR increases with a much clearer mosaic and less noisy than the raw data.

Fig. 12 shows another example of the improvement that the mosaic can bring with respect to the raw data. In this case, the dataset

was collected by a Reson Seabat 7128 mounted onto the ROV Latis from University of Limerick during the ANT'12 trial near La Spezia, Italy (see Fig. 11). This sonar has a much better resolution than BlueView (2.5 cm), lower levels of noise and a higher range (typically in our experiments up to 100 m). Even if the raw data is better than the one provided by BlueView sonars, in some cases objects are not very well defined (close to the centre of Fig. 12a) Instead, in the mosaic shown in Fig. 12b it is clearly identifiable on the left. This figure is also a good example of the improvement in SNR as the mosaic diminishes visibly the noise level of the raw data. This is the kind of improvement that allows a better target recognition in Mine Countermeasures missions.

Comparing to the previous data set, the raw image quality is superior to the one obtained of BlueView data, which is reflected also in the mosaics obtained. Fig. 13 shows two mosaics obtained by processing Reson data. These mosaics represent the same area after 36 and 132 frames. Two conclusions can be drawn from this example. First, comparing to any of the mosaics obtained with BlueView, the mosaic's quality is higher due to the lower noise present in the raw data. However, even in the presence of lower noise levels, the mosaic still improves the definition of objects and shadows present in the scene as it can be easily seen looking at the right side of each mosaic. The difference between Fig. 13a and b is not considerable but some objects are brighter and shadows are better defined in the mosaic built with a higher number of frames (Fig. 13b). It is worth it to note that, although this sonar produces higher resolution images of bigger dimensions, the algorithm still works in real-time up to 4 Hz. To show that feature, a video is attached to this document (MosaicSonar.avi).

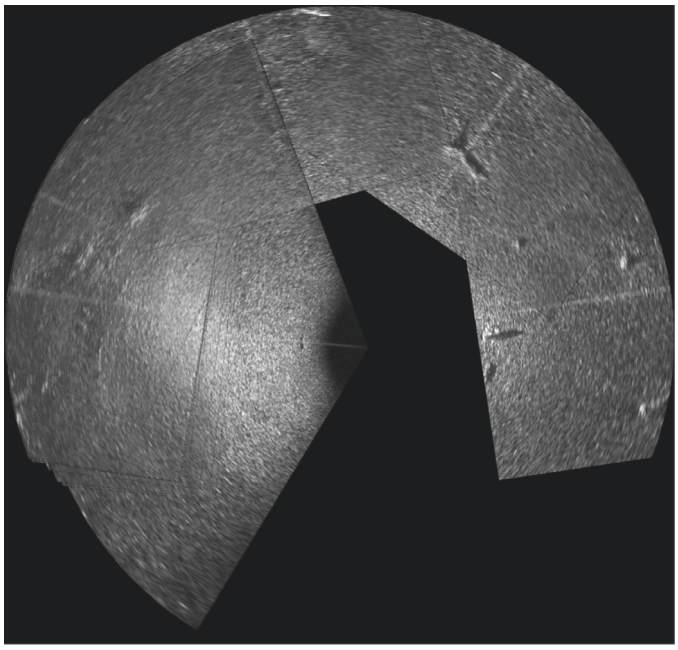
5.2. Mosaicing

A full mosaic obtained in a typical ATR mission is presented in this section to exemplify the kind of result obtained. Fig. 14 shows a mosaic of an area covered during an ATR mission during the ANT'12 trial also organized by CMRE. This mosaic was built from 2511 frames collected with the Reson Seabat 7128 while the ROV was navigating in a cross-pattern mission. The range of the sonar was set to 100 *m* and the area is around 3000 m^2. Differences in insonification are noticed in the mosaic but less than in the example of Fig. 1 due to the higher quality of the sonar. The simple blending technique is not able to produce a fully smooth mosaic. Regarding the purpose of the mission (ATR), in the centre of the mosaic, one can observe a cylindrical shape target. Note that the shadow is much better defined than in BlueView data of Fig. 10. This reflects on the ATR provided that a shadow detector is applied. Even without the shadow detector, ATR has better results on mosaiced data than in raw data. The next section presents quantitative results supporting this fact.

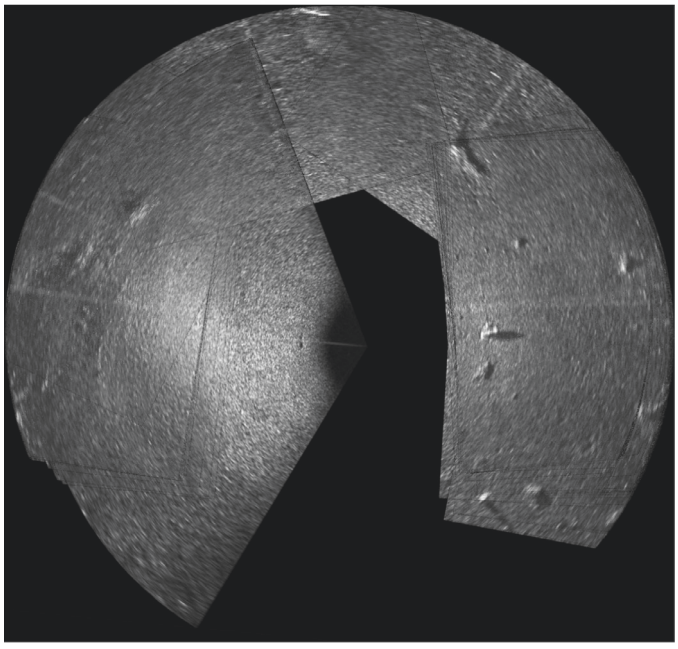
5.3. Automatic target recognition

As mentioned before, one of the best test benches to try the mosaicing output is the Automatic Target Recognition (ATR) in the context of Mine Countermeasures. For the full details on the ATR system applied to FLS data, please refer to the original algorithm (Galceran et al., 2012) and the updated version running on mosaics (Ferreira et al., 2014) which details the improvements of the ATR. Few changes were needed to run the ATR with mosaics instead of raw data. The biggest difference is that instead of processing every single frame, the ATR module now waits for a partial mosaic. Depending on the kind of mission, different criteria are used to decide when to produce partial mosaics. These partial mosaics are then published and can be used in the ATR module.

Several tests were conducted on the integration of the mosaicing algorithm with the ATR module both on collected data and on live sonar. The results shown here are from Multinational AutoNomous Experiment (MANEX'13) as these ones were obtained with



(a) A mosaic built after 36 frames.



(b) The same area after 132 frames.

Fig. 13. Mosaics built with data from Reson SeaBat 5128 mounted onboard the ROV Latis.

live sonars. MANEX'13 was conducted in late October 2013 off the coast of Elba Island, Italy. The sonar is the same as for the Marciana Marina data set. However, a new ASV is used – Gulliver (modified 5.7 m Sea Robotics USV), a catamaran also from CMRE which is slightly bigger than Gemellina. Fig. 15 shows the ASV Gulliver at sea. In Mine Countermeasures (MCM), typical missions executed for target reacquisition are: circling around the bottom target, keeping a distance to that target or cross-pattern missions as seen in Fig. 14. Initial information regarding the position of the target is given by an AUV with an onboard SAS system as described in Section 2.2. Fig. 16 presents a correct target detection in a mosaic even if the raw data is noisy with an interference coming from the echosounder. This interference is translated in several blobs to the left and right of the

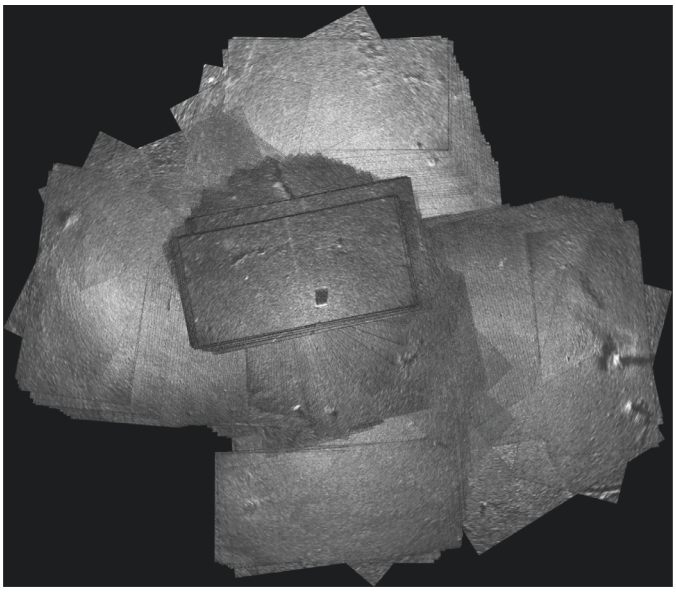


Fig. 14. Mosaic of a cross-pattern mission in an area of 50×60 m^2 approximately.

Fig. 15. The ASV Gulliver from NATO STO CMRE at sea.

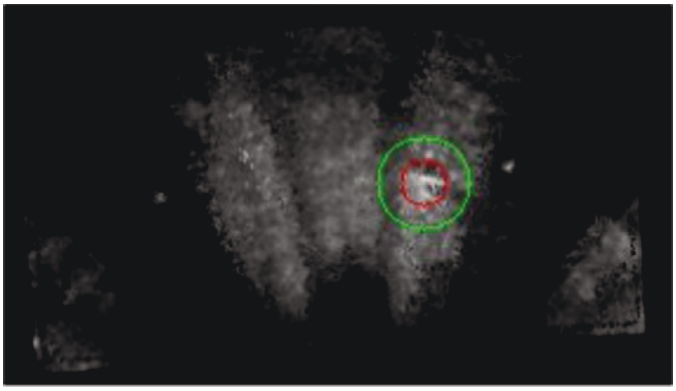


Fig. 16. An example of a correct detection in a mosaic with the presence of noise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Correct and false detections [%], their ratio and total correct detections.

	Correct [%]	False [%]	Ratio	Total correct
Raw data	40.2%	23.4%	1.71	255
Mosaic data	38.4%	15.3%	2.5	15

Table 2
Correct and false detections [%], their ratio and total correct detections for different publishing rates.

	Correct [%]	False [%]	Ratio	Total correct
Raw data	40.2%	23.4%	1.71	255
Mosaic data (20)	38.4%	15.3%	2.5	15
Mosaic data (10)	29.5%	11.2%	2.63	21
Mosaic data (5)	39.7%	14.6%	2.71	60

target. The target (circled in red) is correctly identified and none of the echosounder blobs becomes a false positive. The echosounder interferences are smoothed by the mosaicing process increasing the SNR and allowing a correct recognition. As described in Section 2.1, the echosounder is one of the noise sources that degrades the performance of an ATR system. Thus, the example shown in Fig. 16 is a good example of the worst case scenario in what respects disturbances in the raw data. For the same scenario with such acoustic noise, applying ATR on each raw data frame would result in nearly all false positive detections any time one of the six echosounder signatures appears near the target area.

Fig. 16 can give a qualitative idea of real data used during the trial. However, in order to assess the impact of mosaicing the raw data on the target detection and recognition algorithm, there is the need of establishing quantitative results such as false positive (false detection) and true positive (correct detection) rates. Table 1 shows the results for a data set obtained during a circular mission. Percentages of correct and false detections are computed based on the total number of frames analysed. As discussed in Section 2.2, for circular missions where the target and shadow shape change, a mosaic of the full 360° is not useful due to the artifacts. Instead, partial mosaics limited either by a maximum heading span or number of frames give better results. In this mission, the mosaicing algorithm was publishing a partial mosaic as input to the detection algorithm each 20 frames. This implies that the number of frames analysed by the ATR is 20 times less than if raw data is used. There is a trade-off between the number of frames (and thus the covered area) and the rate of detections. To increase the number of detections, smaller mosaics built from a smaller number of frames should be used with the limit case of one frame for the raw data.

The results of Table 1 show that running ATR on mosaiced data gives a higher ratio of correct detections over false positives. The percentage of correct detections is slightly smaller for the mosaic data than for the raw data but very close to the raw data's rate. Instead, the false positive ratio decreases considerably due to the elimination of noise coming from echosounder harmonics mainly. This allows the mosaic data to have a much better performance in terms of ratio between correct detections and false positives. Nonetheless, the number of detections is small because mosaics are published only each 20 frames. Therefore, the trade-off between number of frames (covered area) and number of detections was further analysed. To a smaller number of frames corresponds a smaller covered area for a given mosaic. The goal of this analysis is to understand the influence of the mosaic publishing rate parameter. Table 2 shows the true (correct) detection and false positive rates and the ratio between them for different publishing rates, namely publishing a mosaic each 5, 10 and 20 frames. These numbers are in parenthesis in each row of the table. The same table includes the results obtained with the raw data for the same data set.

It is easy to conclude that the ratio between correct detections and false positives increases with the mosaic publishing rate. If a mosaic is produced and published each five frames, the number of correct detections increases to 60. Comparing with the raw data (255 detections), this number of correct detections is higher in proportional terms. The best case for the percentage of correct detections is when a mosaic is published each five frames. It does not correspond to the minimum false positives but it has a lower false positive rate than the raw data (14.6% against 23.4%) achieving the highest ratio between correct detections and false positives. The lowest false positive is achieved for the case where a mosaic is published each 10 frames. However, in this case, the percentage of correct detections is also lower than for other publishing rates giving a low total number of correct detections. Regardless of this reduction in correct detections as well as in false positives for this particular case, the ratio between correct and false detections increases with a higher publishing rate and it is substantial higher than for the raw data. For any publishing rate, the ratio between true positives and false positives is always higher than for the raw data. These results show that mosaicing FLS data can improve the ATR of man-manned targets by decreasing the noise level and eliminating several false positives without compromising the True Positive (detection) rate. Although the total number of detections is proportionally smaller than what can be obtained with the raw data, it is still enough for the successive steps of clustering and detections refining. Moreover, it is better to have less detections but a higher number of correct ones than many detections at an expense of a higher number of false positives. By decreasing the number of false positives, mosaic data diminishes their weight on the clustering phase as the ratio between True positives and false positives increases substantially. Consequently, the final result is better as less outliers contribute to the cluster.

5.4. Moving target

Another application of these mosaics, also related to ATR and Mine Countermeasures, is the detection of a moving target. In this case, the mosaics were not used in real-time but they were produced in real-time during the mission. Their integration in the MCM mission is in the planned future work. The data presented in this section was collected using a BlueView P450-45 FLS (45° FOV, 450 kHz operating frequency) during Breaking the Surface (BtS) 2013, the 5th International Interdisciplinary Field Training of Marine Robotics and Applications in the island of Murter, Croatia.[1] The sonar was mounted on a pole fixed to the pier. However, the same results can be extrapolated to the case where the sonar is mounted in a ASV as it was described in Section 2.2. Recalling the application scenario, in this particular setup, besides a target on the sea bottom, there is a small Unmanned Underwater Vehicle (UUV) moving towards that target. This scenario is the successive phase in a MCM scenario after the target detection of the previous section. Mosaicing this quasi-static scene can give the mission history in a very simple and intuitive manner. The name quasi-static comes from the fact that only the UUV is moving versus the target while the ASV (or here the pier) does not move. The mission history can be easily analysed by looking to partial mosaics without the need of replaying the whole sonar file frame by frame.

As proven in Section 5.1, details that cannot be seen, or are hardly seen, in the raw data, are easily detected looking at the produced mosaics. In particular for this application, the UUV missing the target and going forward over the bottom target shadow can be readily identified. In the raw data, this is a very difficult situation to detect in real-time as the shadow prevails. In the mosaic, the UUV position can be observed during the whole mission in an easy manner. Fig. 17 represents a mosaic built from data collected during BtS 2013 with the sonar mounted on a pier. It is an example of the aforementioned situation. In this mission, the UUV missed the target and ended up in the shadow region of the bottom target. This was immediately detected by looking at this mosaic at the end of the mission. A further extension to this data analysis is the use of background subtraction. If background subtraction based on the mosaic is performed, this situation could be prevented. This is because the bottom target position is constant and the only moving object is the UUV as most of the noise gets filtered in the mosaicing process. In this case, the real-time capability becomes important again as the detection of this faulty situation in real-time would allow the correction of the UUV trajectory and the success of the mission. Future work on this is described in the following section.

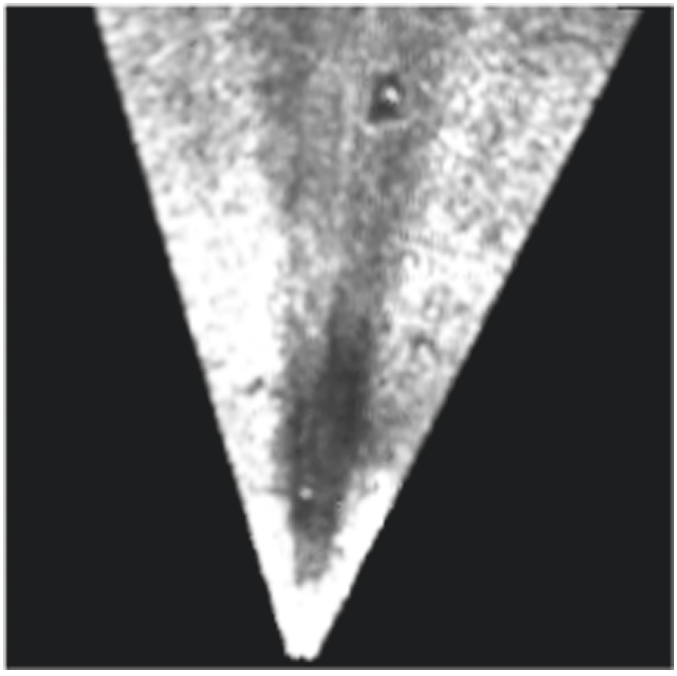


Fig. 17. In the top of the picture, an UUV is depicted in the shadow area of the bottom target.

5.5. Data reduction

Finally, a side effect of no small importance derived from mosaicing FLS data is the reduction in data size compared with the raw data. This is especially important when mosaicing higher resolution acoustic cameras where there is a considerable improvement in terms of computational efficiency. For instance, for the Reson sonar, a raw frame saved in a .png format occupies 1.2MB of space. This sonar was mounted onboard a ROV. For this kind of vehicles, it is common to have high bandwidth fiber optics data link that could eventually ensure the transmission of all raw data in real-time to the mothership. However, if such a sensor is mounted on an AUV that has to transmit the data to an operator via an acoustic link, the amount of raw data created by the sonar is unbearable especially if one takes into account that the maximum ping rate of the Reson sonar is 50 pings per second. Although some optical solutions nowadays start to be interesting with Commercial-of-the-Shelf (COTS) optical modems reaching up to 10 Mbps for a range of 40 m according to Campagnaro, Favaro, Casari, and Zorzi (2014), those are neither very common nor support long distances operations. One option is to compress the raw data, but for a high ping rate the issue still occurs. On the other hand, a mosaic is a collection of several frames but its size does not increase linearly with the number of frames. Although it covers a bigger area it does

[1] http://bts.fer.hr.

Table 3
Comparison between raw data size and mosaic data size.

	Area [m^2]	# frames	Mission time [min]	Total raw data size [MB]	Mosaic data size [KB]	Ratio raw/mosaic
Small	3650	141	9	170	400	425
Large	75,000	8827	42	4400	3200	1375

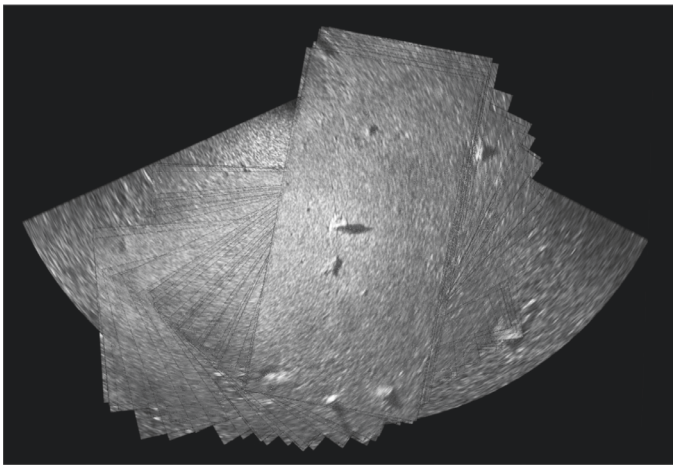


Fig. 18. Mosaic of a scene with several objects present.

not have a much bigger size allowing for efficiency gains. For instance, one situation at which the mosaic is especially useful is if the robot is almost stopped. In this case, sending all the frames would be sending repeated data. Instead, the mosaic can better represent the situation and has the dimension of about one frame. To maintain the real-time capability, the mosaic is built at half resolution of the original data. This does not constitute an issue and it is enough to represent the scene as the raw data has a very high resolution.

To exemplify the data reduction allowed by mosaicing a given scene, Fig. 18 shows a mosaic of an area of interest with several targets during ANT'12 with Reson sonar onboard ROV Latis. It was built based on 141 pings and the area covered is around 3500 m^2. Using the .png format as for the raw data, the mosaic size is around 400KB. Compressing it in a .jpeg format with a 85% compression factor, its size decreases to 200 KB without compromising the mosaic quality. Note that the compression factor is high, decreasing it can lead to further size reduction. If all the 141 frames were sent, that would correspond to around 170MB of raw data without compression. Instead, to send a mosaic of 141 frames shown in Fig. 18, 200 KB are enough. This is a much smaller and especially bearable quantity of data to be sent through an acoustic link.

Sending this compressed mosaic through the acoustic link can be very beneficial to several applications including but not limited to Mine Countermeasures. If the mosaic is built and sent in real-time to a control station, a panoramic view of the area can be assessed by a human operator either for search and rescue operations or human-in-the-loop ATR. For instance, in mine search or reacquisition missions, the mission planned can be adjusted in real-time taking into account the mosaic of the area covered up to that moment. When a mosaic is produced, a very fast and crude ATR algorithm can give a count of the objects present in the area. If this count is higher than a certain threshold, then, the compressed mosaic is sent to the operator on a surface vessel. The operator can evaluate the mosaic, search for the objects of interest and request a more detailed mosaic of the scene around a particular object or group of objects. Along with the mosaic, the coordinates of the vehicle (either in a local or global coordinate system) are available. Thus, the user can click and/or zoom the area generating commands to the robot. If an object is of high interest, then a GO BACK command targeting the object's location can be sent via the acoustic link interrupting the pre-programmed search pattern and defining a new one. After the second look is performed, the robot continues its planned mission.

The fact that mosaicing allows for better human-robot interaction with underwater robots is a key point. By analysing relevant data, a human can interact with the underwater robot by changing the mission and/or parameters such as tilt angle of the sonar, range, etc. This can also diminish operation time and cost as the operator can avoid repassing in areas of no interest and focus on the potential interesting ones.

In order to summarise the computational efficiency gain provided by mosaicing the sonar data, Table 3 shows quantitative data about two of the data sets used. In the case of the Reson sonar, the dataset from Fig. 18 is in analysis while for the BlueView sonar, the Marciana Marina full dataset presented in Ferreira et al. (2014) is the chosen one. Table 3 presents the area covered, the corresponding number of frames, mission time, total raw data size, mosaic data size and ratio between these two for both data sets. The total amount of raw data for a big data set such as the Marciana Marina is in the order of gigabytes while the whole mosaic is around 3MB. Taking into account that partial mosaics can be sent instead of the full one, the order of magnitude of the data to be sent through the acoustic channels becomes kilobytes which is very acceptable. Instead, the gigabytes needed to send all the raw data are completely out of what is possible nowadays.

The improvement in terms of data size is easy to observe by looking at the ratio between the total amount of raw data and the mosaic data. While for the Reson data set, this ratio is 425, for the BlueView case it raises to 1375. For bigger data sets the data size reduction is more accentuated. Taking into account that the first frame appears fully in the mosaic, the accentuated data size reduction is expected. The weight of the dimension of the initial frame in the mosaic is bigger in a smaller data set as the one collected with the Reson sonar. Optimisation techniques can be applied regarding the size of the mosaic to be sent as the acoustic channel has a limited bandwidth. Besides compressing the mosaic, it has to be decided when to send partial mosaics of a given size. The optimal size will be a function of the communication link, total area to be covered, AUV speed, environment, etc. On the operator side, the interface can also include the reconstruction of the global mosaic based on the partial mosaics received from the underwater robot.

6. Conclusion and future work

Several conclusions can be drawn from this work. As mentioned in the abstract, mosaicing FLS can be very useful in the context of Mine Countermeasures for several reasons. First, for the increase of SNR that reflects on a decrease on the false positives rate and better ATR performance. This is allowed because the algorithm works in real-time. It is worthwhile to note that aspect and the fact that it can be applied to different scenarios and applications as shown in other works. A few algorithms are focused on mosaicing in real-time especially in the acoustic domain. Only very recently (after our work was preliminary presented at conferences), Hurtos, Nagappa, Palomeras, and Salvi (2014a) have shown one of the first algorithms working fully in real-time. The first results were obtained for a small area ($17 \times 8 m^2$) and close to the bottom (altitude of 3 m) but look promising

when compared to the offline approach from the same authors. This work is not integrated in a mission that can use the results in real-time as in our case. Therefore, the contribution to the field is relevant and hopefully will stimulate other groups of research to work in the same direction. However, further developments and extensions can and should be performed.

For instance, in featureless environments, image registration does not work properly and GPS data was used to estimate the displacement between frames. Although this can be enough for the purpose of ATR and in small areas, dead reckoning has always some drift. Correlation techniques are used to correct the drift, but correlation may fail in completely featureless environments. Correlation in the frequency domain might be more robust in such cases. Therefore, further work in this direction should focus on first trying to identify that an environment is featureless and then apply different techniques depending on the kind of environment.

Secondly, with respect to the ATR application, there are several possible improvements although more related with the detection itself and not with the mosaicing. Namely, a shape analysis that takes into account the longer axis and smaller axis of a cylinder instead of the horizontal and vertical dimensions in order to improve the detection when the cylinder is rotated shall be investigated. Currently, this technique is being developed but the smaller axis estimate is too noisy in the data obtained with the BlueView sonar. Tests should focus on data collected by the higher resolution Reson SeaBat and then try to transpose the results to the BlueView data.

Other future work has to do with the case of a moving robot in the context of ATR mentioned in Section 5.4. In this case, the mosaic of a quasi-static scene can be used for background subtraction as mentioned before. Subtracting the background in a scene where only the robot is moving, and forgetting the obvious noise, the only blob of the expected size and dimensions always appearing in the image will be the moving robot which makes the work of the detector and tracker much easier. Moreover, as the bottom target is part of the background and thus subtracted, it is now possible to detect the exact moment when the moving robot goes over the target.

The previous paragraph implied that the scene is quasi-static and that the ASV is merely moving and keeping its position. This may not be true but what was said above can be extended to the case where the ASV is moving. This is because the position of the bottom target is known in GPS coordinates and can be translated to the image coordinate system at any time and then subtracted. This extension is useful also in the case where an AUV is used instead of an ASV. The AUVs considered are generally non-holonomic and cannot maintain their position while the moving robot goes to the bottom target. The main point is that as long as the bottom target is on the sonar FOV, the same background extraction principle can be applied.

Finally, the algorithm should be tested in other applications like ship-hull inspection, chain inspection and search and rescue operations. In all of these applications, time can be a crucial factor. With respect to inspections, saving time means saving money. However, when used in search and rescue operations, the benefit is even more important as it can be used for damage assessment and victim recovery.

Acknowledgments

This work was performed under the 2013 CMRE CPOW, and supported by the Centre's Visiting Researcher Programme. Research supported in part by the Fundação para a Ciência e Tecnologia (FCT), Portugal with the PhD Grant SFRH/BD/72024/2010 and by the FP7 IP MORPH Project number 288704. The authors would like to thank Professor Shahriar Negahdaripour (from University of Miami) for the fruitful discussions and Alberto Grati and Dr. Stefano Fioravanti (from CMRE) for their effort during the testing phase. In addition, the authors thank Dr. Warren Fox, ANMCM Programme Manager at CMRE, for his comments during the internal paper review process.

References

AcousticView (2013). Clear Sight software for sonar image enhancement and mosaicking. http://www.acousticview.com/Sonar.htm.

Aykin, M. D., & Negahdaripour, S. (2013). Forward-look 2-D sonar image formation and 3-D reconstruction. In *Proceedings of 2013 MTS/IEEE Oceans, San Diego* (pp. 1–10).

Beaujean, P.-P. J., Brisson, L. N., & Negahdaripour, S. (2011). High-resolution imaging sonar and video technologies for detection and classification of underwater munitions. *Marine Technology Society Journal, 45*(6), 62–74. doi:10.4031/MTSJ.45.6.6.

Benjamin, M. R., Schmidt, H., Newman, P. M., & Leonard, J. J. (2010). Nested autonomy for unmanned marine vehicles with MOOS-IvP. *Journal of Field Robotics, 27*(6), 834–875.

Campagnaro, F., Favaro, F., Casari, P., & Zorzi, M. (2014). On the feasibility of fully wireless remote control for underwater vehicles. In *Proceedings of the 48th Asilomar conference on signals, systems and computers, 2014* (pp. 33–38). doi:10.1109/ACSSC.2014.7094391.

Djapic, V., Nad, D., Ferri, G., Omerdic, E., Dooly, G., Toal, D., & Vukic, Z. (2013). Novel method for underwater navigation aiding using a companion underwater robot as a guiding platforms. In *Proceedings of 2013 MTS/IEEE Oceans, Bergen* (pp. 1–10). doi:10.1109/OCEANS-Bergen.2013.6608153.

Dolbec, M. R. (2007). *Velocity estimation using forward looking sonar*. Naval Postgraduate School, Ph.d. thesis.

Ferreira, F., Djapic, V., & Caccia, M. (2015). Real-time mosaicing of large scale areas with forward looking sonar. *IFAC-Papers on Line, 48*(2), 32–37 ISSN 2405-8963, http://dx.doi.org/10.1016/j.ifacol.2015.06.006.

Ferreira, F., Djapic, V., Micheli, M., & Caccia, M. (2014). Improving automatic target recognition with forward looking sonar mosaics. In *Proceedings of IFAC world congress: 19* (pp. 3382–3387). Cape Town. doi:10.3182/20140824-6-ZA-1003.01485.

Ferreira, F., Veruggio, G., Caccia, M., & Bruzzone, G. (2012). Real-time optical SLAM-based mosaicking for unmanned underwater vehicles. *Intelligent Service Robotics, 5*(1), 55–71.

Galceran, E., Djapic, V., Carreras, M., & Williams, D. P. (2012). A real-time underwater object detection algorithm for multi-beam forward looking sonar. In F. L. Pereira (Ed.), *Proceedings of navigation, guidance and control of underwater vehicles (NGCUV) 2012: 3* (pp. 306–311). Porto.

Hurtos, N., Cuf', X., Petillot, Y., & Salvi, J. (2012). Fourier-based registrations for two-dimensional forward-looking sonar image mosaicing. In *Proceedings of IEEE/RSJ international conference on intelligent robots and systems (IROS), 2012* (pp. 5298–5305). doi:10.1109/IROS.2012.6385813.

Hurtos, N., Cufi, X., & Salvi, J. (2013). A novel blending technique for two-dimensional forward-looking sonar mosaicing. In *Proceedings of 2013 MTS/IEEE Oceans, San Diego* (pp. 1–7).

Hurtos, N., Nagappa, S., Palomeras, N., & Salvi, J. (2014a). Real-time mosaicing with two-dimensional forward-looking sonar. In *Proceedings of 2014 IEEE international conference on robotics and automation (ICRA)* (pp. 601–606). IEEE. doi:10.1109/ICRA.2014.6906916.

Hurtos, N., Palomeras, N., Carrera, A., Carreras, M., Bechlioulis, C. P., Karras, G. C., … Kyriakopoulos, K. (2014b). Sonar-based chain following using an autonomous underwater vehicle. In *Proceedings of 2014 IEEE/RSJ international conference on intelligent robots and systems* (pp. 1978–1983). IEEE. doi:10.1109/IROS.2014.6942825.

Hurtós, N., Ribas, D., Cufí, X., Petillot, Y., & Salvi, J. (2015). Fourier-based registration for robust forward-looking sonar mosaicing in low-visibility underwater environments. *Journal of Field Robotics, 32*(1), 123–151. doi:10.1002/rob.21516.

Johannsson, H., Kaess, M., Englot, B., Hover, F., & Leonard, J. J. (2010). Imaging sonar-aided navigation for autonomous underwater harbor surveillance. In *Proceedings of 2010 IEEE/RSJ international conference on intelligent robots and systems* (pp. 4396–4403). IEEE. doi:10.1109/IROS.2010.5650831.

Karabchevsky, S. (2011). *Real-time underwater obstacle detection using forward looking sonar and FPGA*. Ben-Gurion University of the Negev, Ph.D. thesis.

Kim, K., Neretti, N., & Intrator, N. (2005). Mosaicing of acoustic camera images. *IEE Proceedings Radar, Sonar and Navigation, 152*(4), 263–270. doi:10.1049/ip-rsn:20045015.

Kim, K., Neretti, N., & Intrator, N. (2008). MAP fusion method for superresolution of images with locally varying pixel quality. *International Journal of Imaging Systems and Technology, 18*(4), 242–250. doi:10.1002/ima.20137.

Miskovic, N., Djapic, V., Nad, D., & Vukic, Z. (2011). Multibeam sonar-based navigation of small UUVs for MCM purposes. In S. Bittanti, A. Cenedese, & S. Zampieri (Eds.), *Proceedings of the 18th IFAC world congress: 18* (pp. 14754–14759). Milan: IFAC.

Negahdaripour, S., Aykin, M. D., & Sinnarajah, S. (2011). Dynamic scene analysis and mosaicing of benthic habitats by FS sonar imaging – issues and complexities. In *Proceedings of Oceans 2011* (pp. 1–7). Waikoloa, HI.

Negahdaripour, S., & Xu, X. (2002). Mosaic-based positioning and improved motion-estimation methods for automatic navigation of submersible vehicles. *IEEE Journal of Oceanic Engineering, 27*(1), 79–99.

Newman, P. (2009). Introduction to programming with MOOS. technical report, Oxford Robotics Research Group.

Petillot, Y., Ruiz, I. T., & Lane, D. M. (2001). Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar. *IEEE Journal of Oceanic Engineering, 26*(2), 240–251.

Pizarro, O., Eustice, R. M., & Singh, H. (2009). Large area 3-D reconstructions from underwater optical surveys. *IEEE Journal of Oceanic Engineering, 34*(2), 150–169.

Reed, S., Petillot, Y., & Bell, J. (2004). Automated approach to classification of mine-like objects in side-scan sonar using highlight and shadow information. *IEE Proceedings Radar, Sonar and Navigation, 151*, 4856.

Richmond, K., & Rock, S. (2007). An operational real-time large-scale visual mosaicking and navigation system. *Sea Technology*, (3), 10–13.

SAMM (2014). SAMM stand alone-mosaicking module for forward-looking sonar. http://www.oicinc.com/samm.html.

Singh, H., Roman, C., Pizarro, O., Eustice, R., & Can, A. (2007). Towards high-resolution imaging from underwater vehicles. *The International Journal of Robotics Research, 26*(1), 55–74. doi:10.1177/0278364907074473.

SonarWiz (2013). SonarWiz5 interfaces and file formats. http://www.chesapeaketech.com/docs/SonarWiz5_Interfaces_and_FileFormats.pdf.

Thomas, B., Iv, B. R., & Reed IV, T. B. (2011). Mind the gap! Foward-looking sonar fills in missing data from Nadir 'Holiday'. *Sea Technology, 52*(6), 25–27.

Williams, D. P., & Groen, J. (2011). A fast physics-based, environmentally adaptive underwater object detection algorithm. In *Proceedings of annual conference and exposition on ocean science, Oceans 2011 IEEE, Spain* (pp. 1–7). IEEE. doi:10.1109/Oceans-Spain.2011.6003424.

Yong, E. W. (2011). *Investigation of mosaicing techniques for forward looking sonar*. Heriot-Watt University, University of Girona, Universite de Bourgogne, Master's thesis.

Fausto Ferreira received an Integrated Master Degree in Electrical and Computer Engineering from Instituto Superior Técnico, Technical University of Lisbon in 2008. He defended his PhD in April 2015 in the University of Genoa, Italy. From November 2008 to November 2010, he was a Marie Curie Early Stage Researcher in the EU project FREEsubNET at the National Research Council of Italy (CNR), IEEIT. During his PhD, he had a J-1 Short-Term Courtesy Appointment at University of Miami funded by the Office of Naval Research Global. Currently, he is a Scientist at NATO STO Centre for Maritime Research and Experimentation. Dr. Ferreira has been the Deputy Technical Director of the Student Autonomous Underwater Competition-Europe (SAUC-E) for the 2014 and 2015 editions and of the euRathlon 2014 and 2015 competitions. He has over 30 peer-reviewed publications including a patent.

Vladimir Djapic received the B.S. and M.S. degrees from the University of California at San Diego, in 2000 and 2001, and the Ph.D. degree from the University of California at Riverside, Riverside, in 2009, all in electrical engineering. He returned to the Unmanned Maritime Laboratory in Space and Naval Warfare Systems Center Pacific in San Diego in 2014 where he is a Chief Scientist and a lead Principal Investigator (PI) for projects that utilize Maritime Autonomous Systems (air, surface, and subsurface). From 2008 to 2013 he worked at Center for Maritime Research and Experimentation (CMRE), former NATO Undersea Research Centre (NURC), La Spezia, Italy, and served as a Scientist-in-charge for 5 major NATO sea trials. From 2002 to 2007, he worked at Space and Naval Warfare Systems Center Pacific in San Diego. Dr. Djapic has served as Technical Director of Student Autonomous Underwater Competition-Europe (SAUC-E) since 2010 and since 2013 as a PI for euRathlon and Robocademy projects. He has over 50 publications at prestigious international journals and conferences.

Michele Micheli received a master degree from the University of Pisa in 2002. He is with NATO STO Centre for Maritime Research and Experimentation (CMRE) since 2006 as "Principal Scientific Assistant" in the former Reconnaissance Surveillance and Networks Department. Since then has always worked in the field of sonar signal processing and on the integration of new algorithms into several types of Autonomous Underwater Vehicles (AUV).

Massimo Caccia (MSc 1991) is the Director of CNR-ISSIA. He is member of the Steering Committee of the EU FP7 project CADDY, and he is CNR Principal Investigator of several research projects such as EU FP7 MORPH, EU FP7 CART, EU FP7 MINOAS and in the national PRIN project MARIS. He was international expert in the EU project CURE. He is member of the IFAC Technical Committee 7.2 Marine Systems, of the Board of Directors of the Ligurian District of Marine Technologies and of the Technical Scientific Committee on innovation in shipyards of the Italian Ministry of Infrastructures and Transport. He is author of 8 book chapters, and more than 160 international journal and conference papers.

Document Data Sheet

Security Classification		Project No.
Document Serial No.	**Date of Issue**	**Total Pages**
CMRE-PR-2019-142	June 2019	15 pp.

Author(s)

Fausto Ferreira, Vladimir Djapic, Michele Micheli, Massimo Caccia

Title

Forward looking sonar mosaicing for mine countermeasures

Abstract

Forward looking sonars (FLS) are nowadays popular for many different applications. In particular, they can be used for Automatic Target Recognition (ATR) in the context of Mine Countermeasures. Currently, ATR techniques are applied to raw data which generates many false positives and the need for human supervision. Mosaicing FLS data increases target contrast and thus reduces false positive rate. Moreover, it implies a considerable data size reduction which is important if one thinks of exchange of data in real time through an acoustic channel with very limited bandwidth. Results of applying a real-time mosaicing algorithm to FLS data generated during Mine Countermeasures missions are shown and discussed thoroughly in this article.

Keywords

Mosaicing, forward looking sonar, automatic target recognition, mine countermeasures

Issuing Organization

NATO Science and Technology Organization
Centre for Maritime Research and Experimentation

Viale San Bartolomeo 400, 19126 La Spezia, Italy

[From N. America:
STO CMRE
Unit 31318, Box 19, APO AE 09613-1318]

Tel: +39 0187 527 361
Fax:+39 0187 527 700

E-mail: library@cmre.nato.int