

An Efficient Non-uniformity Correction Technique for Side-Scan Sonar Imagery

Adrian Galdran^{1,2}, Alberto Isasi¹, Mohammed Al-Rawi³, Jonathan Rodriguez³,
Joaquim Bastos⁴, Fredrik Elmgren⁵, and Marc Pinto⁶

¹ Tecnalia Research and Innovation, Derio, Spain, e-mail: alberto.isasi@tecnalia.com

² INESC TEC Porto, Portugal, e-mail: adrian.galdran@inesctec.pt

³ Departamento de Eletrónica, Telecomunicações e Informática (DETI),
Universidade de Aveiro, Aveiro 3810-193, Portugal, e-mails: {al-rawi, jonathan}@ua.pt

⁴ Instituto de Telecomunicações - Pólo de Aveiro, Aveiro 3810-193, Portugal, e-mail: jbastos@av.it.pt

⁵ DeepVision AB, Linköping, Sweden, e-mail: fredrik@deepvision.se

⁶ ECA Robotics, La Garde, France, e-mail: mpi@eca.fr

Abstract—Mapping the seabed represents a fundamental task for many applications. A key technology for that goal is Side-Scan Sonar (SSS) imaging, which offers a large operating range and high resolutions. However, SSS often suffers from echo decay due to water absorption, producing undesired intensity non-uniformities in the image. We propose here a new inhomogeneity correction technique for SSS imagery that exploits two-dimensional information to estimate and remove this non-uniformity. Our approach achieves results similar or better than other recent techniques, and it enjoys a great computational efficiency, being a good candidate for a real-time implementation.

Index Terms—Side Scan Sonar Normalization, Echo Decay, Intensity Inhomogeneity Correction, Autonomous Underwater Vehicle, Seabed Mapping

I. INTRODUCTION

Surveying the seabed represents a task of relevance in many areas, *e.g.* environmental, archaeological, or defense applications [1]–[3]. The predominant underwater sensor modality are acoustics, since its operating range outperforms that of electromagnetic waves, with current Side-Scan Sonar (SSS) providing high resolution images in large ranges of capture.

However, underwater sonar imaging is a challenging task for several reasons. The range and orientation of the sonar head relative to the seabed, together with water characteristics, heavily influences the image acquisition process [4]. Furthermore, the sound waves emitted from the sonar’s towfish, after reflection, need to travel back through a highly absorbing medium. The strength of the received echo is highly dependent on the distance traversed by the sound wave, as well as the angle at which it hits the seabed (grazing angle). As a consequence, a phenomenon known as echo decay tends to appear on the acquired scan. Visually, the echo decay effect appears as an undesired low-frequency brightness gradient similar to the one shown in Figure 1.

This undesired effect due to echo decay may impact the performance of subsequent computer vision tasks, such as object detection [1], registration of different maps of the same area [5], or fusion of those maps [6]. These and many other applications can greatly benefit from a standardizing

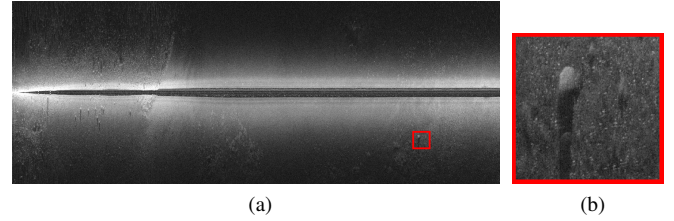


Fig. 1. (a) SSS depicting an inhomogeneous intensity brightness due to echo decay. (b) Close-up detail of a natural landmark. We can observe the undesired effect echo decay may have on automatic seabed analysis algorithms.

technique, capable of correcting the echo decay degradation. For this reason, much research has been dedicated to solve this problem. An early approach was presented in [7], where the authors proposed a pipeline of image processing techniques, designed for the correction of radiometric distortions on SSS degradation due to multiple effects, *e.g.* slant range geometry, speckle noise, multiple returns, or echo decay, also referred to as across-track power drop-off decay. Noise removal has been also addressed in [8] in the along track direction, for the purpose of sonar image mosaics. The noise removal step was shown to improve results of subsequent unsupervised seabed classification algorithms. In [9], noise was also removed in the along track direction, based on the assumption that adjacent pings should contain a similar amount of backscattered energy. In the same work, an heuristic brightness normalization technique is also proposed based on the same principle: for each ping, the average energy of the 20 adjacent pings is computed, and the energy of the ping is adjusted to this averaged profile. A more complex model was proposed in [10], where a sound propagation model that considers also the seabed reflectivity and the wave incidence angle is proposed. This model is related to the received echo through a Lambertian model of the seabed surface, which can be inverted to remove uneven echo decay. A more comprehensive review of different approaches for sonar image pre-processing can be found in [11].

In this work, we propose a new technique for correcting the non-uniformity present in SSS due to echo decay. The method

makes efficient use of visual information from different sub-areas of the scan, locally normalizing the intensity in each region to retrieve a more regular image. Experimental results show a good behavior of our technique, outperforming other recent approaches both qualitatively and quantitatively.

II. A NEW TECHNIQUE FOR ILLUMINATION COMPENSATION ON SSS

To model low-frequency intensity variations in SSS, here we consider a degradation scheme composed of a luminosity and a contrast drift, that generate the observed scan as follows:

$$\tilde{S}(x, y) = C(x, y)S(x, y) + L(x, y), \quad (1)$$

where $\tilde{S}(x, y)$ is the observed non-uniform sonar scan, $S(x, y)$ is the uniform scan we intend to estimate, and $C(x, y)$, $L(x, y)$ represent respectively the contrast and luminosity distortion.

Note that the total dynamic range of raw data acquired by standard sonars exceeds that of common display devices and standard images. Commonly, sonar manufacturers implement a proprietary dynamic range compression with the objective of reducing the dynamic range. This is useful to enable the application of computer vision and image analysis techniques. To simplify our exposition, we assume that a simple logarithmic dynamic range compression is applied to compress the received acoustic signal.

To solve the same model (1) in the context of retinal imaging, in [12] the authors proposed to measure $C(x, y)$ and $L(x, y)$ from a decomposition of the undegraded image $S(x, y)$ into background and foreground. Following this approach, we consider that the acquired scan can be represented as an additive combination of a background image S_{bg} that captures low-frequency content, and a foreground image S_{fg} containing objects, landmarks, and other high-frequency components on the seabed:

$$S(x, y) = S_{bg}(x, y) + S_{fg}(x, y). \quad (2)$$

Following [12], to model the background information present around each pixel (x, y) in a SSS, we consider a normal distribution centered on it:

$$S_{bg}(x, y) \sim \mathcal{N}(\mu_{bg}, \sigma_{bg}). \quad (3)$$

Inserting eq. (2) into eq. (1), we have that:

$$\begin{aligned} \tilde{S}(x, y) &= C(x, y) (S_{bg}(x, y) + S_{fg}(x, y)) + L(x, y) \\ &= C(x, y)S_{bg}(x, y) + C(x, y)S_{fg}(x, y) + L(x, y) \end{aligned} \quad (4)$$

Now, we consider the action of the above model only on pixels lying on the background, i.e. $(x, y) \in B$. On these pixels, the foreground image has no content, and $\tilde{S}_{fg}(x, y) = 0$. Hence, we can write:

$$\tilde{S}(x, y) = C(x, y)S_{bg}(x, y) + L(x, y) \quad \forall (x, y) \in B. \quad (5)$$

Since we have assumed that the background follows a normal distribution, we deduce that for background pixels the mean

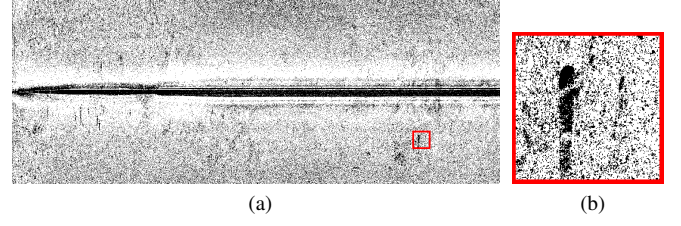


Fig. 2. (a) Binary background image corresponding to the SSS depicted in Figure 1(a). Background was computed according to eq. (7) for $\tau = 1$. (b) The same natural landmark shown in Figure 1(b). Note how most of the landmark and its acoustic shadow are considered foreground.

and standard deviation of that distribution approximate the luminosity and contrast drift:

$$S(x, y) \sim \mathcal{N}(C(x, y), L(x, y)), \quad \forall (x, y) \in B. \quad (6)$$

Thus, we can estimate luminosity and contrast distortions from background pixels by looking into local average and standard deviation values on the background. In practice, we are assuming the echo decay effect affects mostly the background image, which contains the low-frequency information, and for that reason, echo decay can be characterized processing only background pixels.

To isolate background pixels from the initial sonar scan, we again follow [12]. In this case, we need to assume that in local neighborhoods $\Omega(x, y)$ around pixels (x, y) , both luminosity and contrast drifts remain constant. This is meaningful, since both are slowly-varying phenomena. Second, we consider also that for each local neighborhood, there is a sufficient amount of background pixels such that the background intensity distribution can be reliably estimated. Again this is a reasonable assumption, since natural landmarks and other objects in the seabed represent a small subset of the entire sonar image. Finally, we also assume that background pixels have intensities relatively different to foreground pixels (otherwise, the separation into two classes would not be feasible in terms of scan intensities only).

Under the above hypothesis, we need to locally estimate the parameters defining the normal distribution given by eq. (6). The simplest choice is to employ the local mean and standard deviation on each neighborhood. With this, the background can be easily extracted: if a pixel (x, y) belongs to the background B , its intensity will be close to the mean intensity in $\Omega(x, y)$.

However, in order to consider also the local variation in intensities, instead of applying an Euclidean distance, we rather compute a Mahalanobis distance of each pixel intensity from the local model given by $(\mu_{\Omega}, \sigma_{\Omega})$, i.e., a pixel (x, y) belongs to the background if the following holds:

$$S_{bg}(x, y) = \left| \frac{S(x, y) - \mu_{\Omega}}{\sigma_{\Omega}} \right| < \tau, \quad (7)$$

being τ a threshold, which is a parameter of the algorithm.

Figure (3a) shows the background image of the scan in Figure (2a), and we also show the same close-up zoom of the region containing a natural landmark in Figure (3b). The background image was constructed with $\tau = 1$. This value

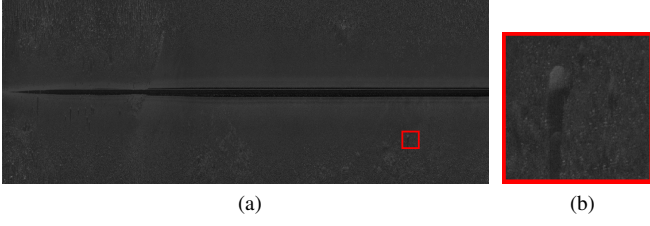


Fig. 3. (a) The SSS depicted in Figure (1a), after correction with the proposed method for echo decay removal. (b) The same natural landmark shown in Figure (1b), with a clearly reduced echo decay effect.

ensures that $\sim 68\%$ of the pixels will be considered as background, due to the normal distribution of the luminosity.

Once we have an estimate of the background component of the input sonar $S(x, y)$, we can easily estimate contrast and luminosity drifts. According to eq. (6), these can be identified with the local mean and standard deviation, but extracted only from pixels in S_{bg} . Thus, we only need to compute again local mean and standard deviations from local neighbourhoods, this time on $S_{bg}(x, y)$, and assign them to our estimates $\hat{L}(x, y)$, $\hat{C}(x, y)$ of $L(x, y)$ and $C(x, y)$. These are finally used to retrieve a corrected sonar image, by directly inverting eq. (1):

$$S(x, y) \approx \frac{\tilde{S}(x, y) - \hat{L}(x, y)}{\hat{C}(x, y)}. \quad (8)$$

The output of eq. (8) contains negative and positive values. We simply apply an affine transformation to back-project every intensity to the $[0, 1]$ range. The result of the outlined technique, applied to the SSS in Figure 1, for a square neighborhood with side 200 pixels, is shown in Figure 3.

A. Efficient Implementation

Since implementation of eqs. (7) and (8) implies only a subtraction and a division, the main computational load of the method is the calculation of means and standard deviations across local neighborhood on the non-uniform sonar image, which is a large resolution image. These computations need to be performed twice, since the first local estimates allow to build the background image, from where the luminosity and contrast drifts are estimated afterwards by computing again local statistics.

In a brute-force approach, the above procedure could be implemented by looping on every image pixel, and computing means and standard deviation for each of them. Fortunately, both operations can be computed in linear time with the help of integral images and box-filters [13]. This simple idea lies at the core of many modern computer vision algorithms that are intended to perform efficiently, such as the Viola and Jones object detection framework [14].

To apply integral images and box-filters for estimating local means, we consider spatial-neighborhoods $\Omega(x, y)$ of square shape, with radius r , around pixels (x, y) . In this setting, an integral image is built by storing at the (x, y) coordinate sim-

ply the sum of all pixel intensities corresponding to locations above and to the left of (x, y) :

$$I(x, y) = \sum_{i \leq x, j \leq y} S(i, j). \quad (9)$$

In addition, the integral image can be constructed in a single pass over the initial scan S . If the computations proceed by scanning the image from the leftmost top pixel, values in adjacent pixels depend only on a few previously computed values:

$$I(x, y) = S(x, y) - I(x-1, y-1) + I(x, y-1) + I(x-1, y). \quad (10)$$

This process is exemplified in Figure 4. Once the integral image $I(x, y)$ has been built out of the initial scan $S(x, y)$, it serves as a look-up table. Thus, the evaluation of mean intensities across $\Omega(x, y)$ requires only four operations. That is, if we consider a square neighborhood of corners $C_0 = (x_0, y_0)$, $C_1 = (x_1, y_0)$, $C_2 = (x_0, y_1)$, and $C_3 = (x_1, y_1)$, the sum of $S(x, y)$ over the rectangle defined by C_0 , C_1 , C_2 , and C_3 is given by:

$$\sum_{\substack{x_0 < x \leq x_1 \\ y_0 < y \leq y_1}} S(x, y) = I(C_2) + I(C_0) - I(C_1) - I(C_3). \quad (11)$$

Moreover, this calculation (referred to as box-filtering) is independent of the radius r , *i.e.* we obtain a constant computational time.

After normalization by division over the total amount of pixels in $\Omega(x, y)$, we obtain the local mean with just one addition, two subtractions, and one division per pixel, independently of the size of $\Omega(x, y)$.

To derive local variances, we recall its definition:

$$\sigma^2(x, y) = \frac{1}{n} \sum_{i=1}^n (S(i, j) - \mu(x, y))^2, \quad (12)$$

where n is the number of pixels in $\Omega(x, y)$. Hence, we need to build both the integral image corresponding to the input scan I_S and the square of it, I_{S^2} . With that, we can then easily estimate local variances:

$$\sigma^2(x, y) = \frac{1}{n} \left(I_{S^2}(x, y) - \frac{I_S^2(x, y)}{n} \right). \quad (13)$$

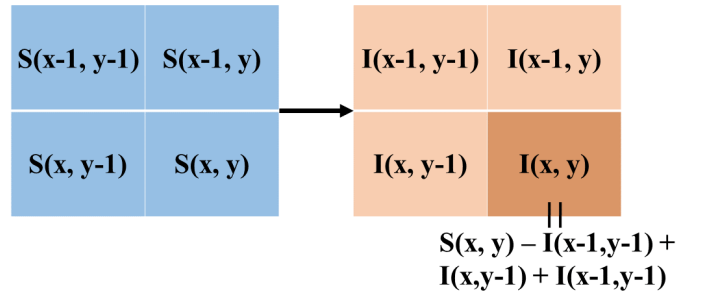


Fig. 4. Computation of an Integral Image value. Note that the value at the (x, y) position only depends on adjacent values.

With this approach, both the local mean and local variance calculations are $\mathcal{O}(N)$, independently of the size of the considered spatial neighborhoods.

Moreover, a straightforward strategy to further reduce the complexity of the above process has been suggested in [17]. For that, we first subsample by bicubic interpolation the initial sonar scan by a factor of s . Then, every box filter - eq. (11) - is computed on this low-resolution scan. After local estimates for the mean and the standard deviation are computed at the low resolution level, we again upsample them, and the remaining computations take place at high resolution. Since these two operations are the main computational burden of the proposed technique, the efficiency is highly influenced by this speed-up: the entire procedure complexity decreases from $\mathcal{O}(N)$ to $\mathcal{O}(N/s^2)$, see [17].

It is worth noting that a similar strategy was followed in [12] to increase the performance of their algorithm. However, in our case efficiency concerns are more critical, since the SSS we deal with can be of a considerably large resolution, greater than 2000×6000 pixels.

III. EXPERIMENTAL RESULTS

We provide now qualitative and quantitative evaluation of the proposed technique, comparing it to the DCP [15] and the MIRA [16] methods. Our data consists of a set of five SSS recorded in Lake Vättern (Sweden) by DeepVision. The images were recorded with a DeepEye 680D SSS (Chirp sonar, 680kHz). Combining left and right scans, each sonar image has more than 6000 pings with each ping having 2000 samples.

A. Qualitative Analysis

The effect of applying our technique can be qualitatively appreciated in Figure (5a–5d). In Figure (5e), a sub-region of the entire SSS is shown. We see that the inhomogeneity that dominates the original scan is removed to some extent by the DCP, and more effectively by MIRA and our method. Note that MIRA employs one ping to estimate the echo decay on each side of the scan. Even with such little information, the method is able to remove non-uniformity from the image. However, the computational cost of MIRA is high. Moreover, if we want to improve its estimate by using more pings, that cost will raise to important times. On the contrary, our technique makes use of information coming from different areas of the image, and it has the potential to be implemented in real time.

B. Quantitative Evaluation

For quantitative evaluation, we employ the Sonar Image Quality Evaluation Metric (SIQEM), proposed in [16]. This metric measures the global homogeneity of a SSS by comparing the mean homogeneity of different subregions on it.

To obtain a global estimate of the homogeneity of a SSS, SIQEM analyzes the difference of echo values between adjacent regions. For that, the scan is divided into a set of patches covering the whole scanned seabed. Then, for each patch (i, j) the mean intensity is computed, and denote by $\mu_{i,j}$. Finally, to estimate the quality of an image patch, SIQEM compares its

TABLE I
SIQEM RESULTS FOR EACH SSS PROCESSED BY EACH TECHNIQUE.

Method	Side-Scan Sonars				
	W1	W2	W3	W4	W5
None	0.115	0.120	0.117	0.109	0.137
DCP	0.129	0.143	0.142	0.125	0.157
MIRA	0.076	0.087	0.080	0.076	0.107
Ours	0.042	0.051	0.045	0.043	0.057

average intensity with that of neighboring regions following Webers perceptual law:

$$C_{i,j} = \frac{\max(\mu_{i,j} - \mu_{i_k,j_k}, \mu_{i_k,j_k} - \mu_{i,j})}{\mu_{i,j}}, \quad (14)$$

where k indexes patches adjacent to (i, j) . Finally the SIQEM metric is derived from eq. (14) as:

$$SIQEM(S) = \frac{1}{M \cdot N} \sum_{i=1}^M \sum_{j=1}^N C_{m,n}, \quad (15)$$

where the total number of patches is $M \times N$. The lower SIQEM value an image attains, the less inhomogeneity it contains.

Results of computing SIQEM for the set of five SSS is displayed in Table I. We see that the proposed technique clearly outperforms both DCP and MIRA in every SSS, implying that it is able of removing much effectively the observed non-uniformity.

We also measure the impact of the proposed technique by analyzing the across-track energy changes for every ping. Even if the method employs two-dimensional information to normalize intensities, it is expected that this normalization successfully translates to uni-dimensional regions on the SSS. To measure this effect, we compute the coefficient of variation for each ping on each of our five SSS. For a given ping denoted as S_{i,j_0} , for a fixed j_0 , and i varying from the border of the image until the nadir section, this coefficient is defined as the quotient of the standard deviation with respect to the mean.:

$$CV(S_{i,j_0}) = \frac{\sigma(S_{i,j_0})}{\mu(S_{i,j_0})}. \quad (16)$$

In Figure 6 we can see a plot of these values for each of our five SSSs. We can see how the original scan contains a considerable variation in most cases. This variation is reduced by every considered method. However, we can observe that the SSS normalized by the proposed technique contains significantly lower signal variations along pings, demonstrating the benefits of our approach. This can be numerically confirmed in Table II, where we show average of the across-track CVs shown in Figure 6.

IV. CONCLUSION

In this work, a new inhomogeneity correction technique for Side Scan Sonar imagery is proposed. The method makes use

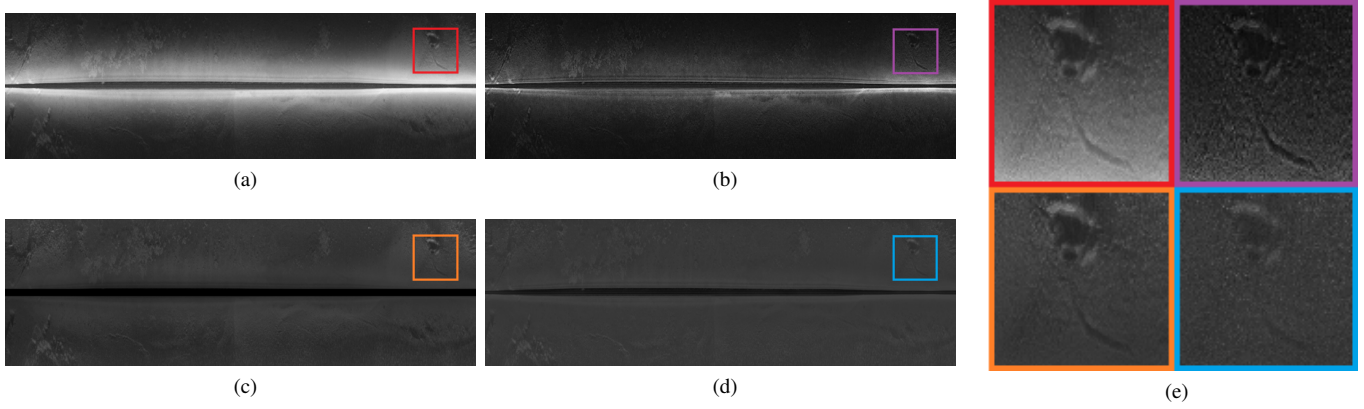


Fig. 5. One of the acquired Side Scan Sonar Images, and its corrected versions. a) Original scan b) Corrected by DCP [15] c) Corrected by MIRA [16] d) Corrected by the proposed technique e) Close-up details showing a landmark in the seabed for each of the images.

TABLE II
AVERAGE CV FOR ACROSS-TRACK PINGS FOR EACH SSS PROCESSED BY EACH TECHNIQUE.

Method	Side-Scan Sonars				
	W1	W2	W3	W4	W5
None	0.692	0.677	0.688	0.715	0.724
DCP	0.827	0.787	0.787	0.854	0.866
MIRA	0.323	0.334	0.331	0.328	0.348
Ours	0.178	0.191	0.183	0.179	0.167

of two-dimensional information to estimate of the presence of non-uniformities and effectively remove them. The proposed technique offers results similar or better than other recent methods, and it enjoys a great computational efficiency, being a good candidate for a real-time implementation.

ACKNOWLEDGMENT

This work is supported by SWARMs EU project (Smart & Networking Underwater Robots in Cooperation Meshes), Grant Agreement n. 662107-SWARMs-ECSEL-2014-1, which is partially supported by the ECSEL JU and the Fundao para a Ciênciã e a Tecnologia (ECSEL/0002/2014 and ECSEL/0003/2014).

REFERENCES

- [1] E. Dura, Y. Zhang, X. Liao, G. J. Dobeck, and L. Carin, "Active learning for detection of mine-like objects in side-scan sonar imagery," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 2, pp. 360–371, Apr. 2005.
- [2] J. J. Schultz, C. A. Healy, K. Parker, and B. Lowers, "Detecting submerged objects: The application of side scan sonar to forensic contexts," *Forensic Science International*, vol. 231, no. 13, pp. 306–316, Sep. 2013.
- [3] M. Reggiannini and O. Salvetti, "Seafloor analysis and understanding for underwater archeology," *Journal of Cultural Heritage*, vol. 24, pp. 147–156, Mar. 2017.
- [4] R. P. Hodges, *Underwater Acoustics: Analysis, Design and Performance of Sonar*. John Wiley & Sons, Jun. 2011.
- [5] N. Hurtos, D. Ribas, X. Cuf, Y. Petillot, and J. Salvi, "Fourier-based Registration for Robust Forward-looking Sonar Mosaicing in Low-visibility Underwater Environments," *Journal of Field Robotics*, vol. 32, no. 1, pp. 123–151, Jan. 2015.
- [6] S. Reed, I. T. Ruiz, C. Capus, and Y. Petillot, "The fusion of large scale classified side-scan sonar image mosaics," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2049–2060, Jul. 2006.
- [7] R. L. Miller, F. S. Dwan, and T. W. C. Hilde, "Image Processing Techniques For The Preprocessing Of Side Scan Sonar Data," in *Geoscience and Remote Sensing Symposium, 1990. IGARSS '90. Remote Sensing Science for the Nineties'.*, 10th Annual International, May 1990, pp. 2281–2281.
- [8] D. Wilken, P. Feldens, T. Wunderlich, and C. Heinrich, "Application of 2d Fourier filtering for elimination of stripe noise in side-scan sonar mosaics," *Geo-Marine Letters*, vol. 32, no. 4, pp. 337–347, Jun. 2012.
- [9] Y.-C. Chang, S.-K. Hsu, and C.-H. Tsai, "Sidescan sonar image processing: correcting brightness variation and patching gaps," *Journal of Marine Science and Technology*, vol. 18, no. 6, pp. 785–789, 2010.
- [10] A. Burguera and G. Oliver, "Intensity correction of Side-Scan Sonar images," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, Sep. 2014, pp. 1–4.
- [11] Y.-S. Shin, Y. Cho, Y. Lee, H.-T. Choi, and A. Kim, "Comparative Study of Sonar Image Processing for Underwater Navigation," *Journal of Ocean Engineering and Technology*, vol. 30, no. 3, pp. 214–220, 2016.
- [12] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis*, vol. 9, no. 3, pp. 179–190, Jun. 2005.
- [13] F. C. Crow, "Summed-area Tables for Texture Mapping," in *Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '84. New York, NY, USA: ACM, 1984, pp. 207–212.
- [14] P. Viola and M. Jones, "Robust Real-time Object Detection," in *International Journal of Computer Vision*, 2001.
- [15] K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
- [16] M. Al-Rawi, A. Galdran, X. Yuan, M. Eckert, J. Martinez, F. Elmgren, B. Curuklu, J. Rodriguez, J. Bastos, and M. Pinto, "Intensity Normalization of Sidescan Sonar Imagery," in *International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oulu (Finland), Dec. 2016.
- [17] K. He and J. Sun, "Fast Guided Filter," *arXiv:1505.00996 [cs]*, May 2015, arXiv: 1505.00996. [Online]. Available: <http://arxiv.org/abs/1505.00996>

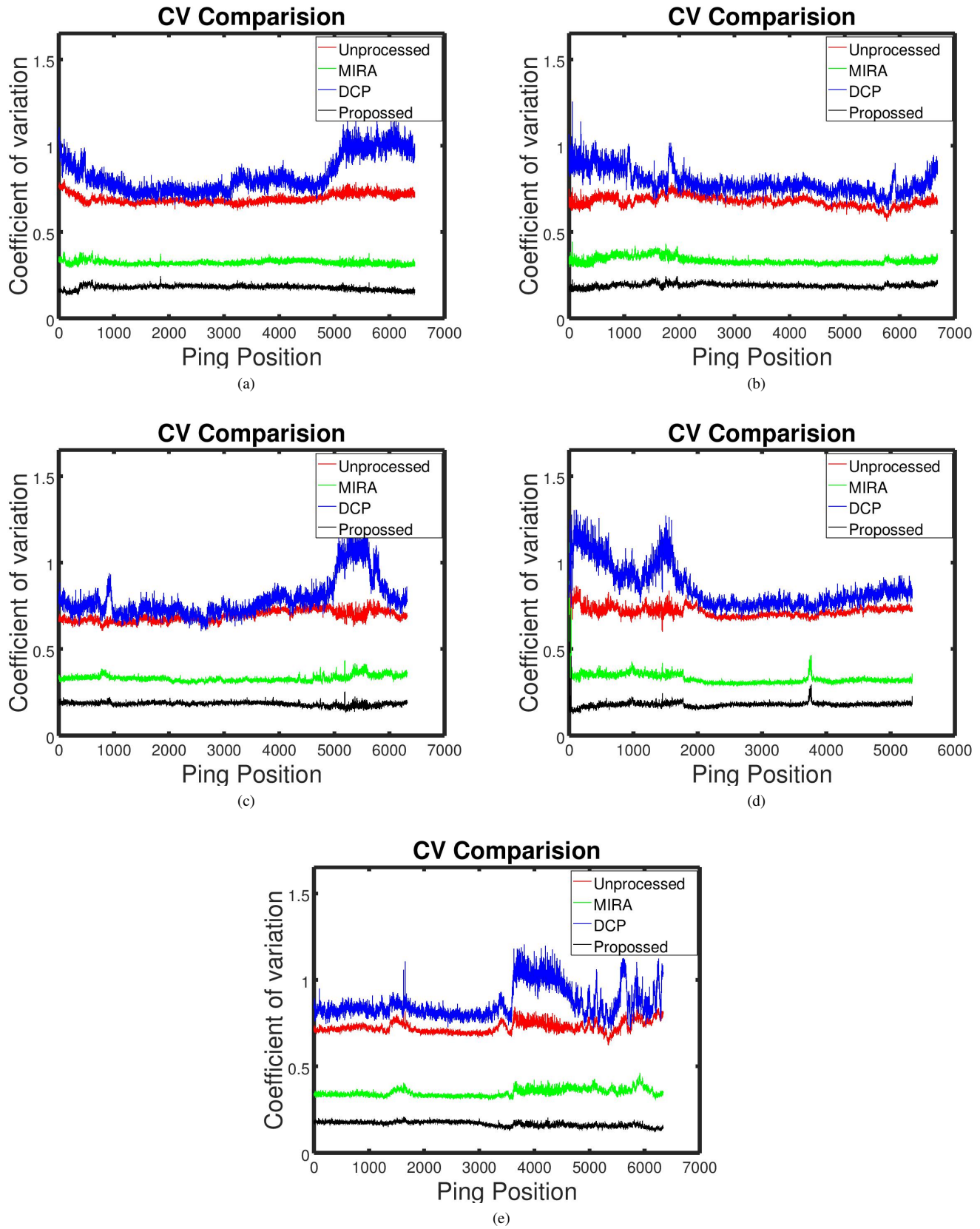


Fig. 6. (a)-(e) Coefficient of variation for each ping, measured across-track. Each plot corresponds to each of the SSS in our dataset. Note that to avoid interference of the nadir region, we consider only one side of all scans.