SCIENCE AND TECHNOLOGY ORGANIZATION
CENTRE FOR MARITIME RESEARCH AND EXPERIMENTATION

Reprint Series CMRE-PR-2019-008

The new muesli complexity metric for mine-hunting difficulty in sonar images

David P. Williams

May 2019

Originally presented at:

2018 OCEANS – MTS/IEEE Kobe Techno-Ocean (OTO)
doi: 10.1109/OCEANSKOBE.2018.8559193

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.

The New Muesli Complexity Metric for Mine-Hunting Difficulty in Sonar Images

David P. Williams
NATO STO
Centre for Maritime Research and Experimentation
La Spezia, Italy
david.williams@cmre.nato.int

***Abstract*—A new image complexity metric has been developed that fuses the concept of lacunarity, a measure of pixel intensity variation, with the notion of spatial information, a quantity that captures edge energy. This new metric, which we call the "muesli" complexity, successfully quantifies the relative difficulty of performing target detection in synthetic aperture sonar (SAS) images. This has been experimentally validated via the results of a human operator study, as well as the results of an object detection algorithm, using a set of over 3000 SAS images collected in diverse environments. In the former assessment method, it has been observed that the subjective human rankings of image difficulty correlate well with the complexity value. In the latter examination approach, it has been observed that the degrees to which false alarms are generated and true targets are missed by the detection algorithm are each proportional to the complexity value of the image.**

***Index Terms*—Image complexity, performance estimation, mine countermeasures (MCM), automatic target recognition (ATR), synthetic aperture sonar (SAS)**

I. Introduction

In terms of mine countermeasures (MCM), quantifying the complexity of an image enables one to predict or estimate performance, or adapt parameters of automatic target recognition (ATR) algorithms. Several different complexity metrics have recently been proposed in the literature for sonar imagery, including approaches based on wavelets [1]–[3] and lacunarity [4]. The purpose of this work is to introduce and evaluate a new complexity metric.

The remainder of this paper is organized as follows. Sec. II introduces the new muesli image complexity metric, while Sec. III details the sonar data on which it is assessed. Methods for establishing ground truth via human operator experiments and detection algorithm quantities are described in Sec. IV. Experimental evaluation of the new metric is presented in Sec. V, before concluding remarks are made in Sec. VI.

II. Image Complexity Metrics

Lacunarity was originally developed as a way to measure spatial structure in binary-valued data [5], but the concept has since been extended to quantify pixel-intensity variation in grayscale imagery [6]. In our formulation, lacunarity is closely related to concepts such as scintillation index [7], the (inverse) shape parameter of a k-distribution [8], and coefficient of variation [9], among others.

The lacunarity of a set of pixels in a grayscale image is the ratio of the variance of the pixel values to the square of the mean of the pixel values. Formally, we define the lacunarity for pixels x_i whose indices are in a set N of cardinality n as

$$L = \frac{\sigma^2}{\mu^2}, \tag{1}$$

where

$$\sigma^2 = \frac{1}{n}\sum_{i\in N}(x_i - \mu)^2 \tag{2}$$

and

$$\mu = \frac{1}{n}\sum_{i\in N} x_i. \tag{3}$$

When the set N corresponds to indices that constitute a rectangular block of pixels, this calculation can be done very quickly using integral images [10]. Repeating this procedure about each pixel in an image effectively creates a "map" of the lacunarity quantifying local pixel-intensity variation.

In contrast, spatial information [11] is a quantity that captures edge energy. Let s_h and s_v represent the result of filtering an image $\mathbf{I}$ with horizontal and vertical Sobel kernels, respectively. The spatial information at each pixel can then be obtained as

$$\mathbf{S}(\mathbf{I}) = \sqrt{s_h^2 + s_v^2}. \tag{4}$$

In this work, however, rather than using the original sonar image $\mathbf{I}$ as the input for calculating the spatial information, we instead use the lacunarity map, $\mathbf{L}(\mathbf{I})$, computed from image $\mathbf{I}$ over a predefined window whose size is related to the objects of interest. (In this work, the window is 0.75 m$\times 0.75$ m.) The mean of this particular spatial information quantity, averaged over all pixels, is then the metric of interest. That is, the new image complexity metric is the mean of the spatial information of the lacunarity of the original sonar image: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I})))$. For the sake of brevity, we coin the term "muesli" to represent this new operator, $\mu(\mathbf{S}(\mathbf{L}(\cdot)))$. (The name alludes to both the operator's mixture nature and its lexicography.)

Since the utility of lacunarity for characterizing the seafloor in sonar imagery was studied in [4], we compare the proposed muesli complexity metric to five closely related potential alternative metrics, including:

- the mean of the lacunarity of the original sonar image, $\mu(\mathbf{L}(\mathbf{I}))$;

TABLE I
DATA SET CHARACTERISTICS

Data Set Code	Name of Sea Experiment	Survey Dates (months / year)	Survey Location	Number of Images
COL2	Colossus 2	4-5 / 2008	Riga/Liepaja, Latvia	272
CAT1	Catharsis 1	3 / 2009	Palmaria, Italy	40
CAT2	Catharsis 2	10 / 2009	Elba, Italy	165
AMI1	AMiCa	5-6 / 2010	Tellaro, Italy	349
ARI1	ARISE'11	5 / 2011	Bonassola, Italy	320
ARI2	ARISE'12	10-11 / 2012	Elba, Italy	326
SPM1	SPMEX	4 / 2013	Cartagena, Spain	255
MAN1	MANEX'13	9-10 / 2013	Elba, Italy	507
MAN2	MANEX'14	9-10 / 2014	Bonassola, Italy	641
NSM1	NSMEX	5 / 2015	Ostend, Belgium	252
TJM1	TJMEX	10 / 2015	Cartagena, Spain	502
ONM1	ONMEX	9 / 2016	Hyères, France	94

- the variance of the lacunarity of the original sonar image, $\sigma^2(\mathbf{L}(\mathbf{I}))$;
- the lacunarity of the lacunarity of the original sonar image, $L(\mathbf{L}(\mathbf{I}))$;
- the variance of the spatial information of the lacunarity of the original sonar image, $\sigma^2(\mathbf{S}(\mathbf{L}(\mathbf{I})))$;
- the lacunarity of the spatial information of the lacunarity of the original sonar image, $L(\mathbf{S}(\mathbf{L}(\mathbf{I})))$.

In each case, the complexity metric associated with a sonar image scene is a scalar.

III. SONAR DATA SETS

To assess the value of the proposed image complexity metric for establishing mine-hunting difficulty in sonar imagery, we utilize a substantial amount of data collected by CMRE's MUSCLE autonomous underwater vehicle (AUV). This experimental, state-of-the-art AUV is a 21-inch diameter vehicle from Bluefin that is equipped with a high-frequency SAS system developed by Thales. The center frequency of the SAS is 300 kHz, and the bandwidth is 60 kHz. The system enables the formation of high-resolution sonar imagery with a theoretical along-track resolution of 2.5 cm, and a theoretical across-track resolution of 1.25 cm, usually out to a range of 150 m. (Example imagery will be shown later.)

Measured SAS data was collected during twelve sea experiments at various geographical sites in the Mediterranean, Baltic, and North Seas between 2008 and 2016. At each site, man-made targets of various realistic mine shapes – cylinders, truncated cones, wedges – had been deployed before MUSCLE AUV data-collection surveys were executed. For the present study, 3723 SAS image scenes were considered, each of which contains at least one target. A summary of the data set characteristics is shown in Table I.

Example images illustrating the diversity of the environments across the various sea experiments are shown in Fig. 1; the values of the proposed muesli complexity metric are also provided.

IV. GROUND TRUTH

To determine which complexity metrics accurately reflect the difficulty of an image for mine-hunting, some notion of ground truth must first be established. Two approaches are used for this purpose: the ratings of several different human operators, and quantities related to the performance of an automated detection algorithm.

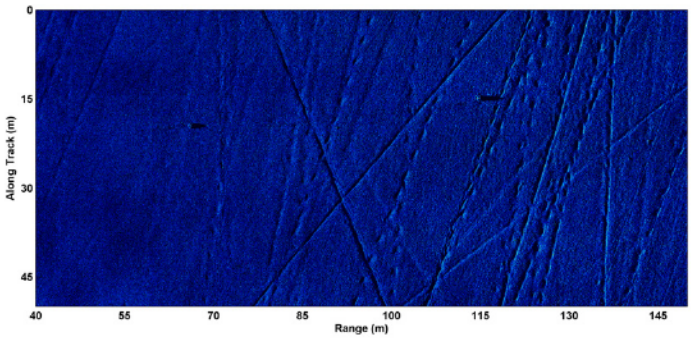
A. Human Operator Ratings

The principal task of an MCM operator is to identify targets in sonar imagery. Experiments were conducted to obtain the *subjective* ratings of humans, acting as operators, regarding the complexity of SAS images from the data set. Specifically, each participant was instructed to rate the complexity of a given SAS image – on a scale of 1 to 5, with 5 being most complex – in terms of how difficult he or she perceived that mine-hunting would be in the image due to the environment. (The participants were not asked to detect targets in the imagery.) The order of the images presented to each human was random.

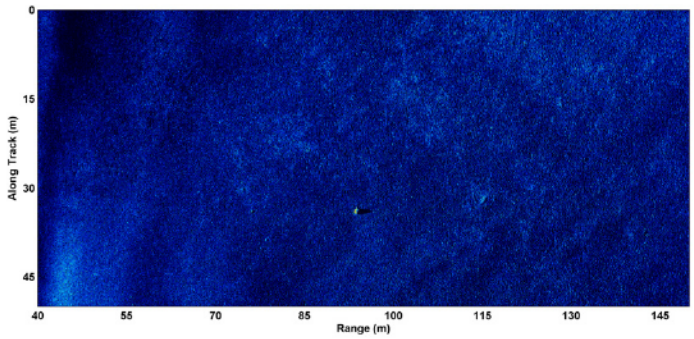
Six humans participated in these operator experiments. Each participant was also asked to self-assess his or her experience level working with sonar imagery, on a scale of 0 (no experience) to 10. The results of this self-assessment, and the number of images that each human rated, are shown in Table II. It can be seen that the participants formed a diverse group in terms of experience.

TABLE II
HUMAN OPERATOR SUMMARY

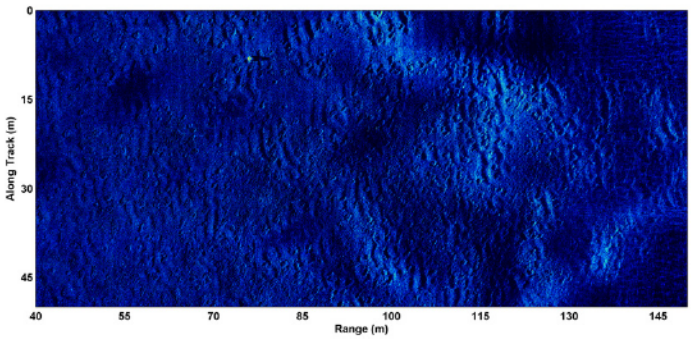
Operator Index	Experience Level	Number of Images Rated
1	9	3723
2	8	3723
3	5	2402
4	3	420
5	1	774
6	0	1611



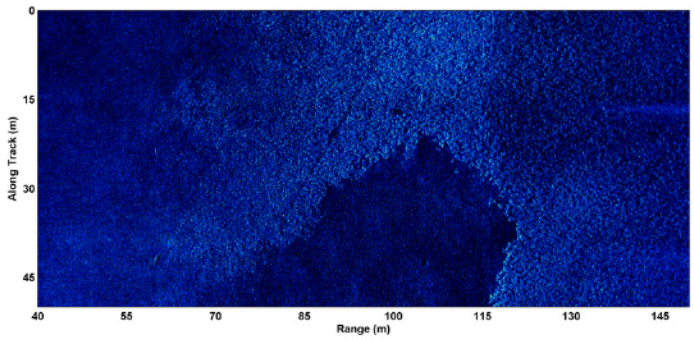
(a) SPM1: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.0713$



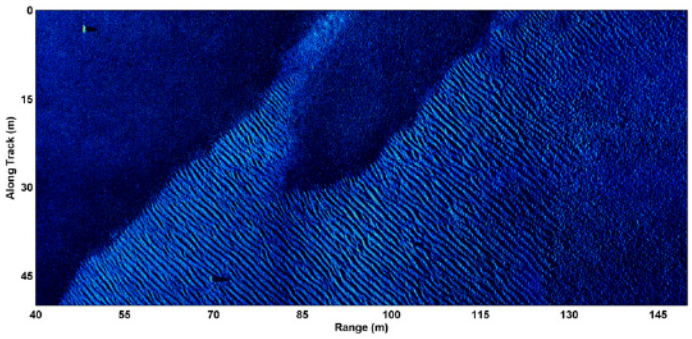
(b) MAN1: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.0835$



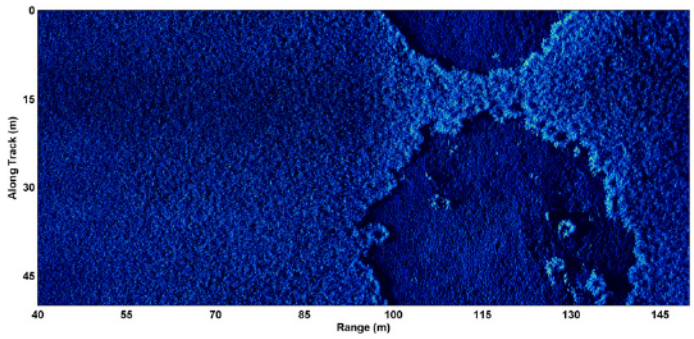
(c) AMI1: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.0917$



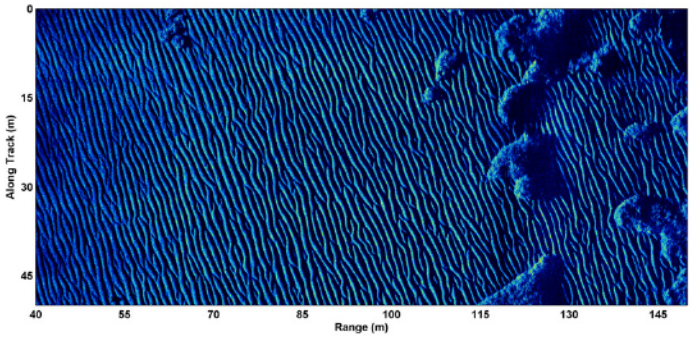
(d) ARI2: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.1743$



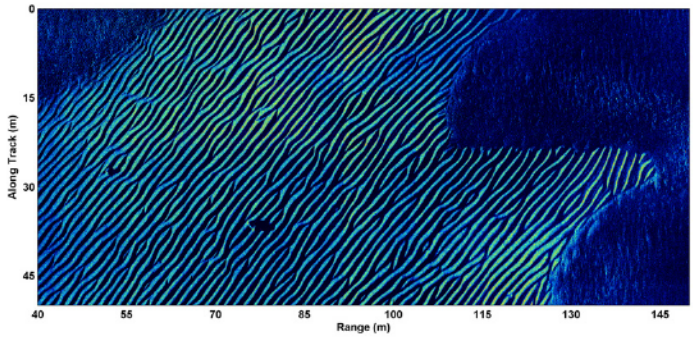
(e) ARI1: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.1787$



(f) CAT2: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.2625$



(g) MAN2: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.2908$



(h) COL2: $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I}))) = 0.3394$

Fig. 1. Example SAS scenes in diverse environments from 8 different sea experiments, indicated in the captions. The value of the proposed muesli complexity metric for each image is also shown.

B. Detection Algorithm Quantities

In addition to the human operator ratings, two other quantities based on the performance of an automated detection algorithm were also used to more *objectively* determine image complexity. The detection algorithm employed, called the Mondrian detector [12], performs a series of tests based on the average pixel intensity in different blocks, or regions, that share a special predefined spatial arrangement in an image. The algorithm was applied to each image in the data set. For a given image, the algorithm generates a set of alarms, each associated with a score $p \in [0, 1]$ that can be viewed as the probability of the alarm being a target.

Mine-hunting may be difficult in an image for two distinct reasons: because a large number of clutter objects (*i.e.,* false alarms) are present, or because the seafloor conditions make the detection of true targets challenging. Therefore, we define the difficulty of performing mine-hunting in a given image via two complementary measures, one related to the propensity of false alarms, and one related to the ability to detect true targets. These two aspects jointly characterize performance.

Specifically, for an image $\mathbf{I}$, we call the former quantity the "false alarm score sum," and define it as

$$\phi(\mathbf{I}) = \sum_{n=1}^{N} p_n, \tag{5}$$

where p_n is the nth false alarm's probability of being a target – *i.e.,* its "score" from the detector – and N is the number of false alarms generated in the image by the detector. This quantity is more informative than a simple count of the number of false alarms, and it allows finer gradations in complexity (*e.g.,* the values are continuous rather than discrete).

The second performance quantity we call the "missed target score sum," and define it as

$$\chi(\mathbf{I}) = 1 - \frac{1}{M} \sum_{m=1}^{M} p_m, \tag{6}$$

where p_m is the mth true target's probability of being a target – *i.e.,* its "score" from the detector – and M is the number of true targets in the image, with the understanding that $p_m = 0$ for any target that goes undetected. It can readily be seen that this quantity expresses the degree to which targets are *not* detected with high probability. For example, when all targets in an image are not detected, $\chi(\mathbf{I}) = 1$. When all targets are detected with very high scores, $\chi(\mathbf{I})$ will be close to 0. This (continuous) quantity is again more informative than a simple (discrete) count of the number or fraction of missed targets.

Thus, for an image deemed to be of high complexity, one would expect that $\phi(\mathbf{I})$ and/or $\chi(\mathbf{I})$ would be large.

C. Validation

The standard (Pearson) correlation measures the strength of the linear relationship between two variables. Given the nature of the ground truth quantities, however, it instead makes more sense to employ the Spearman *rank* correlation, which measures the degree to which the relationship between two variables can be described by a monotonic – not necessarily linear – function. That is, we are interested in determining whether the complexity metrics listed in Sec. II – and particularly the muesli complexity metric – increase as the human operator ranking of complexity increases, and as the false alarm score sum and missed target score sum increase. Thus, the term "correlation" in this work always refers to the Spearman correlation.

First we examine the correlation between the complexity ratings of the different human operators. These pairwise correlations are shown in Table III. (Only the images that were rated by both operators are used in the calculations.) It can be seen that the ratings are rather strongly correlated, indicating general agreement in terms of image complexity across operators. In the table, the highest (non-auto) correlation for each operator, row-wise, is in bold. Interestingly, it can also be observed that the correlation of ratings is typically higher between more experienced operators (*cf.* Table II).

Next, the correlation between a human operator's rating and each of the two detection algorithm quantities are computed, with these results shown in Table IV. It can be seen that stronger agreement exists with the false alarm score sum, ϕ, than with the missed target score sum, χ. This result potentially suggests that the human operators' ratings of image complexity are driven more by the prevalence of false alarm sources in the image than by factors that might obscure a target's detection. Future studies will examine this hypothesis in greater depth.

Having established the reasonableness of the ground truth quantities, we can now proceed to assessing the accuracy of the complexity metrics.

V. Experimental Results

To evaluate the feasibility of the various complexity metrics, the correlation between a given metric and each of the human operator ratings are calculated, with these results shown in Table V. (Also shown is the average correlation between each metric and the operator ratings, when averaged across the six human operators' values.) It can be observed that the proposed muesli complexity metric, $\mu(\mathbf{S}(\mathbf{L}(\mathbf{I})))$, achieves the maximum correlation for each operator (entries in bold).

The environments in the various sea experiments can be markedly different, so we show these correlations broken down by experiment as well, in Fig. 2. Although the trends are generally similar across operators for a given sea experiment, some differences arise (*e.g.,* in CAT1 and SPM1). It will be of interest to examine the characteristics of the images from those sea experiments to better understand the sources of disagreement among operators.

Next, the correlation between a given metric and the two detection quantities are calculated, with these results shown in Table VI. It can be observed that the proposed muesli complexity metric again is strongly correlated to the detection quantities, though $\sigma^2(\mathbf{L}(\mathbf{I}))$ achieves the highest correlation for the false alarm score sum. Of interest is the fact that the correlation with the false alarm score sum is consistently

TABLE III
SPEARMAN CORRELATION BETWEEN HUMAN OPERATOR COMPLEXITY RATINGS

	Operator Index					
Operator Index	1	2	3	4	5	6
1	1.0000	**0.8551**	0.7057	0.7187	0.7014	0.6442
2	**0.8551**	1.0000	0.6787	0.6839	0.6717	0.6538
3	**0.7057**	0.6787	1.0000	0.5922	0.6798	0.6538
4	0.7187	0.6839	0.5922	1.0000	**0.7752**	0.6654
5	0.7014	0.6717	0.6798	**0.7752**	1.0000	0.5541
6	0.6442	0.6538	0.6538	**0.6654**	0.5541	1.0000

TABLE IV
SPEARMAN CORRELATION BETWEEN HUMAN OPERATOR COMPLEXITY RATINGS AND DETECTION ALGORITHM QUANTITIES

Operator Index	Correlation with False Alarm Score Sum	Correlation with Missed Target Score Sum
1	0.6221	0.4298
2	0.6009	0.4461
3	0.4955	0.3874
4	0.5022	0.4137
5	0.5283	0.4380
6	0.4394	0.6243

TABLE V
SPEARMAN CORRELATION BETWEEN COMPLEXITY METRICS AND HUMAN OPERATOR RATINGS

Complexity Metric	Operator Index 1	2	3	4	5	6	Average
$\mu(\mathbf{S}(\mathbf{L}(\mathbf{I})))$	**0.5884**	**0.5293**	**0.4051**	**0.3872**	**0.4560**	**0.4228**	**0.4648**
$\sigma^2(\mathbf{S}(\mathbf{L}(\mathbf{I})))$	0.3912	0.3475	0.2971	0.2278	0.3113	0.2668	0.3070
$L(\mathbf{S}(\mathbf{L}(\mathbf{I})))$	-0.3855	-0.3569	-0.2404	-0.2694	-0.3355	-0.3010	-0.3148
$\mu(\mathbf{L}(\mathbf{I}))$	0.5177	0.4602	0.3268	0.3166	0.3919	0.3653	0.3964
$\sigma^2(\mathbf{L}(\mathbf{I}))$	0.5180	0.4729	0.3908	0.3481	0.4092	0.3823	0.4202
$L(\mathbf{L}(\mathbf{I}))$	0.3219	0.2986	0.2908	0.2597	0.2660	0.2293	0.2777

TABLE VI
SPEARMAN CORRELATION BETWEEN COMPLEXITY METRICS AND DETECTION ALGORITHM QUANTITIES

Complexity Metric	Correlation with False Alarm Score Sum	Correlation with Missed Target Score Sum
$\mu(\mathbf{S}(\mathbf{L}(\mathbf{I})))$	0.5119	**0.2207**
$\sigma^2(\mathbf{S}(\mathbf{L}(\mathbf{I})))$	0.5030	0.0359
$L(\mathbf{S}(\mathbf{L}(\mathbf{I})))$	-0.0988	-0.3449
$\mu(\mathbf{L}(\mathbf{I}))$	0.4080	0.2094
$\sigma^2(\mathbf{L}(\mathbf{I}))$	**0.5743**	0.1377
$L(\mathbf{L}(\mathbf{I}))$	0.5205	-0.0182

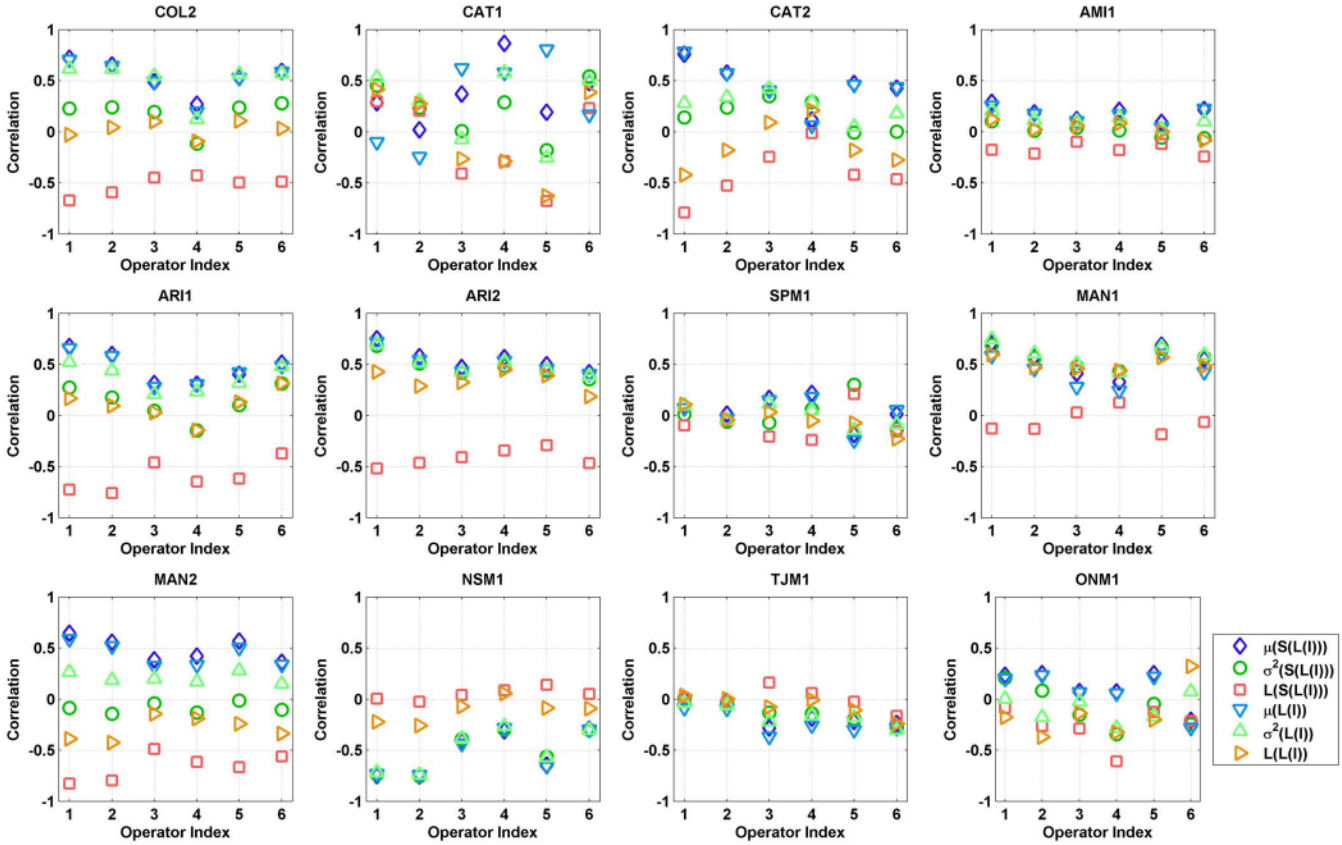


Fig. 2. The correlation of each complexity metric with the human operator ratings, separated by sea experiment.

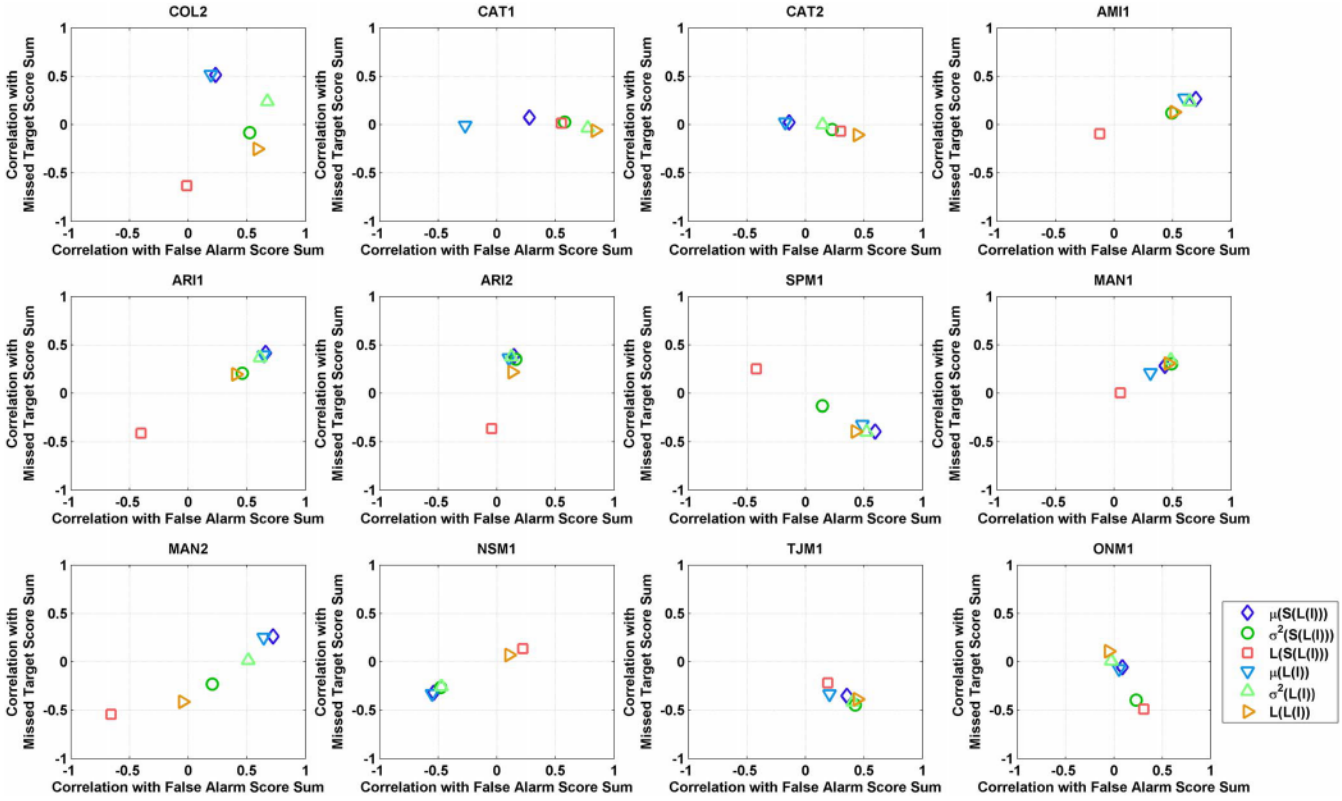


Fig. 3. The correlation of each complexity metric with the false alarm score sum and the missed target score sum, separated by sea experiment.

stronger than the correlation with the missed target score sum. This result may indicate that the complexity metrics are driven mainly by the presence of seafloor features that generate false alarms.

Finally, these correlations are shown as a function of sea experiment in Fig. 3. It will be of interest to examine in the future the factors that create the variations observed. For example, in NSM1, where the complexity metrics did not correlate strongly with the detection quantities, much of the imagery was of poor quality due to currents and uncompensated vehicle motion.

VI. Conclusion

A new complexity metric was developed for establishing the difficulty of mine-hunting in sonar imagery. It was demonstrated that this so-called muesli complexity metric was strongly correlated to human operators' subjective ratings of image complexity, as well as to more objective quantities based on an automated detection algorithm's measures of performance.

Future work will examine the correlation of other complexity metrics from the literature that have been previously developed for sonar imagery. Additional human operator experiments will also be undertaken.

Acknowledgment

This work was supported by the United States Office of Naval Research.

References

[1] O. Daniell, Y. Petillot, and S. Reed, "Unsupervised sea-floor classification for automatic target recognition," in *Proc. International Conference on Underwater Remote Sensing*, 2012.

[2] E. Fakiris, D. Williams, M. Couillard, and W. Fox, "Sea-floor acoustic anisotropy and complexity assessment towards prediction of ATR performance," in *Proc. International Conference and Exhibition on Underwater Acoustics*, 2013, pp. 1277–1284.

[3] M. Geilhufe and Ø. Midtgaard, "Quantifying the complexity in sonar images for MCM performance estimation," in *Proc. International Conference and Exhibition on Underwater Acoustics*, 2014, pp. 1041–1048.

[4] D. Williams, "Fast unsupervised seafloor characterization in sonar imagery using lacunarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6022–6034, 2015.

[5] B. Mandelbrot, *The Fractal Geometry of Nature*. Macmillan, 1983, vol. 173.

[6] R. Plotnick, R. Gardner, W. Hargrove, K. Prestegaard, and M. Perlmutter, "Lacunarity analysis: A general technique for the analysis of spatial patterns," *Physical Review E*, vol. 53, no. 5, pp. 5461–5468, 1996.

[7] E. Fremouw and A. Ishimaru, "Intensity scintillation index and mean apparent radar cross section on monostatic and bistatic paths," *Radio Science*, vol. 27, no. 4, pp. 539–543, 1992.

[8] D. Abraham and A. Lyons, "Novel physical interpretations of k-distributed reverberation," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 4, pp. 800–813, 2002.

[9] S. Aja-Fernández and C. Alberola-López, "On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering," *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2694–2701, 2006.

[10] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[11] H. Yu and S. Winkler, "Image complexity and spatial information," in *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2013, pp. 12–17.

[12] D. Williams, "The Mondrian detection algorithm for sonar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 1091–1102, 2018.

Document Data Sheet

Security Classification		Project No.
Document Serial No.	**Date of Issue**	**Total Pages**
CMRE-PR-2019-008	May 2019	7 pp.

Author(s)

David P. Williams

Title

The new muesli complexity metric for mine-hunting difficulty in sonar images

Abstract

A new image complexity metric has been developed that fuses the concept of lacunarity, a measure of pixel intensity variation, with the notion of spatial information, a quantity that captures edge energy. This new metric, which we call the "muesli" complexity, successfully quantifies the relative difficulty of performing target detection in synthetic aperture sonar (SAS) images. This has been experimentally validated via the results of a human operator study, as well as the results of an object detection algorithm, using a set of over 3000 SAS images collected in diverse environments. In the former assessment method, it has been observed that the subjective human rankings of image difficulty correlate well with the complexity value. In the latter examination approach, it has been observed that the degrees to which false alarms are generated and true targets are missed by the detection algorithm are each proportional to the complexity value of the image.

Keywords

Image complexity, performance estimation, mine countermeasures (MCM), automatic target recognition (ATR), synthetic aperture sonar (SAS)

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