IMPROVEMENT OF AUV-BORNE SEABED MAPPING WITH QUALITY MAPS USING STATISTICAL ANALYSIS

Naveed ISLAM[a], Ahmed NAIT-CHABANE[a], Benoit ZERR[a], Yann DUPAS[b]

[a] Lab-STICC, UMR CNRS 6285, Equipe: Ocean Sensing and Mapping (OSM), ENSTA Bretagne, 2 rue François Verny, 29200, Brest, France,
{naveed.islam, ahmed.nait_chabane, benoit.zerr}@ensta-bretange.fr

[b] SHOM/DO/HOM/DEV, CS 92803 29228 Brest CEDEX 2, France,
Yann.Dupas@shom.fr

Abstract:

Sonar data is commonly affected by noise due to the processing of scatter signals and interference of acoustic waves scattered from the seabed. To overcome this problem and to limit the noise in sonar images, the sonar operator can change the sonar settings (e.g. range, pulse length, modulation, inter-track distance, etc.) to acquire the best possible acoustic data. On board autonomous underwater vehicles (AUV), due to the low bandwidth of the communication with the robot, the real time definition of the best settings by an operator is nearly unfeasible. For these reasons, we have developed an analysis method for automatically assessing the quality of the data. The results of this process are then sent to the AUV planning module which can change the sonar settings (e.g. inter-track distance).

The classical approach is based on the correction of the artefacts related to the wave propagation in water column and the characteristics of the sonar system. This approach requires strong a priori knowledge of the system and the conditions of acquisition of the sonar data.

The main objective of this paper is to propose a statistical measure of quality of the sonar data acquired using AUVs. This statistical measure would be representing a quality map for the input sonar data. As no prior measurement of similarity or dissimilarity of sonar images is given, the decision to whether accept the quality of data as noisy/non-noisy will be based upon statistical hypothesis testing. To accomplish the quality mapping, spectral domain filtering is performed to extract the residual image representing the speckle. Based on Maximum Likelihood (ML) method, parameters are estimated from the data for Rayleigh distribution and its fit is evaluated using Goodness-of-fit (Gof) test. Experimental results show the viability of the proposed approach while mapping the data into quality matrix representing the acceptable regions on sonar data acquired using DAURADE.

Keywords: *Quality measure, Cartography mapping, Noise models, Sonar data, AUV (Autonomous Underwater Vehicle)*

1. INTRODUCTION

The Autonomous Underwater Vehicle (AUV) is becoming a common vehicle to collect acoustic images of the seabed. The sonar mounted on the vehicle sends ultrasonic pulses in a defined direction and records the signal resulting from the interaction on the transmitted pulse and the environment (scattering, reflection, attenuation, etc).

Coherent imaging systems (like synthetic aperture radars, sonar, ultrasound and laser imaging) are commonly affected by multiplicative noise (also known as speckle noise). The speckle noise in these systems is caused by the addition of coherent and random interference in the backscatter signals [1]. A trivial solution to suppress the noise is to directly apply the de-noising low-pass filter, but by directly applying the de-noising filters, relevant information can be lost which significantly affects the later processes of object detection, identification and recognition.

Generally, the speckle noise can be represented by: $D(x,y) = I(x,y) * r(x,y)$, where $D(x,y)$ is the observed sonar image, $I(x,y)$ is the original image and $r(x,y)$ is the multiplicative component of the speckle noise. The de-speckling filters can be categorized into two main groups: Statistical based and Frequency based filters [2]. The known de-speckling filters explained in [3] (like Lee, Frost and Kuan) are statistical based filters which used a-priori statistical models of speckle noise. Frequency domain filters (like wavelet and Fourier) are non-adaptive filters which take all the signal components into consideration and then process the data [4, 5].

All these de-noising filters smooth out the noise while retain the features in the image but cannot provide any qualitative assessment about the data before and after filtering. Since no prior information is given to describe or estimate the quality of sonar image, a sonar image can be describe as a good quality image if the seabed features appear very well just like they would have been observed under the ideal operating conditions of the sonar system [6].

The main objective of this work is: quantitative analyses of the sonar image by providing a quality map of the data into good and bad regions along the track of AUV in different ranges of sonar image. This transformation of quantitative sonar data into quality map would give to the AUV, a better understanding on how the seabed is acquired by the sonar and would ultimately lead to time and costs saving, by optimal real-time mission re-planning.

The rest of the paper is organised as follows. In section 2, the proposed approach for quality mapping is discussed. Experimental results based on the proposed approach over sonar data is given in section 3. Concluding remarks and future prospects are provided in section 4.

2. PROPOSED QUALITY MAPPING APPRAOCH

The information content in the sonar data in terms of quality degrades significantly across the range. It is widely accepted that the noise in the sonar data can be appropriately modelled by the Rayleigh distribution [7]. This can be exploited in the real case scenarios where the data is, acquired through DURADE AUV, having different scale of noise in different ranges of the swath. A hypothesis about the quality of data can be set on the basis of the Rayleigh distribution and the goal of making these inferences can be achieved by testing the hypothesis. The proposed approach can be summarized into the following steps:

- First, the sonar image is partitioned into equal sized windows and each image window is transformed from the spatial domain into frequency domain using Fourier transform.
- Apply the Butterworth high-pass filter in the frequency domain image, which gradually transition from 0 to 1 to keep high frequencies outside a radius and discard the low frequency values inside the radius.
- Calculate the residual image by applying the inverse Fourier transform which contains only the speckle noise representing the high frequency components in the image.
- Estimate the parameters for the Rayleigh distribution from the data using the Maximum likelihood estimation (MLE) method.
- Using hypothesis testing (goodness of fit test), measure how well the Rayleigh distribution fits to the observed residual data.
- Based on the probability of the support in the goodness of fit test, the input sonar data is mapped into a quality matrix.

In the remaining section, the detail of the proposed approach is explained and justified.

A. Spectral Domain Filtering and Analysis:

The amplitude and the phase of the backscattered signal recorded on the sonar transducers are statistically independent. The amplitude is a function of the object reflectivity and the phase is a function of the surface shape. In the domain of image processing, the Discrete Fourier Transform (DFT) is widely used in numerous applications like image analysis, enhancement, filtering, reconstruction and compression [8]. In Fourier transformed image, the low frequencies correspond to the slowly varying information (e.g., homogeneous areas), while the high frequencies correspond to the quickly varying information (e.g., edges). The speckle noise components in the input sonar image mostly belong to the high frequencies in the Fourier space, thus the information content about the speckle can be obtained by filtering the low frequency components in the transformed domain.

The Butterworth filter has the property to gradually supress the frequencies, where the roll-off (sharpness/slope of the transition from the pass-band to the stop-band) is controlled by the filter order [8]. The Butterworth high-pass filter keeps the frequencies outside a radius r_0 and discards those values inside the radius r_0. The high-pass Butterworth filter is given by:

$$H(u,v) = \frac{1}{1+\left(\frac{r_0}{W(u,v)}\right)^{2n}} \tag{1}$$

Where $w(u,v)$ denotes the distance from the centre of the spectrum, r_0 denotes the cut-off frequency which controls the radial size of the filter and n denotes the order of the filter which controls the transition from stop-band to pass-band (i.e. from 0 to 1). A family of filters can be created by varying n to increase or decrease the slope r_0. The Fourier domain image $F(u,v)$ is multiplied with the Butterworth high-pass filter $H(u,v)$ of same size, to produce a filtered image given by:

$$Z(u,v) = F(u,v) * H(u,v) \tag{2}$$

The filtered image $Z(u, v)$ contains the speckle noise components representing the high frequency components in the given sonar image.

Beside the issues related to the configuration of sonar systems, the unresolved problem in any sonar data quality mapping is the lack of substantial prior information or ground truth sonar image which can be used to make a comparison for quality assessment. In the case of limited prior information related to the noise, a quality mapping can be achieved. The only information about the noise can be obtained from residual image $R_t(x, y)$, which is computed by taking the inverse DFT of the filtered image $Z(u, v)$. The residual image only contain information about the speckle component in the sonar image as the low frequency component representing the homogeneous areas in the sonar image have gradually been removed. This operation is very important for understanding the behavior of the noise, as the noise remain the same even in different regions of the seabed's, therefore only their distribution can be used to find their model.

B. Quality Mapping: Model Estimation and Goodness-of-Fit Test

The Rayleigh distribution is widely used to study the speckle noise in coherent imaging systems. The probability density function of the Rayleigh distribution is given by [9]:

$$s(x; \sigma) = \frac{x}{\sigma^2} e^{\frac{-x^2}{(2\sigma^2)}}, x \geq 0, \tag{3}$$

where $\sigma > 0$, is the scale parameter of the distribution. The cumulative distribution function is given by:

$$S(x) = 1 - e^{\frac{-x^2}{(2\sigma^2)}}, \ for \ \ x \in [0, \infty) \tag{4}$$

In order to understand the behaviour of noise in the residual image, it is very important to model its distribution. For modelling any distribution, we need to find or estimate the parameters of the assumed distribution. The objective is to identify the good parameters for the Rayleigh distribution that is mostly likely to have generated the speckle vector R_t. The two most commonly used methods for parameter estimations are the least-square estimation (LSE) and maximum likelihood estimation (MLE) [10]. We choose to estimate the model parameters using MLE because it is more useful in hypotheses testing or constructing confidence intervals and inference in statistics.

Goodness-of-fit(Gof) techniques examine how well a sample of data agrees with a given distribution as its population [10]. Some important Gof tests are; (i) chi-square tests, (ii) moment ratio, (iii) correlation based tests, (iv) empirical distribution function based tests. The Gof procedure defines a test statistics, which measure the distance between the hypothesis and the observed data, and then calculate the probability of obtaining the data, assuming the hypothesis is true. The smaller probability will indicate poor fit, while high probability corresponds to the good fit. Due to the intrinsic restrictions (due to the size of data, distribution type etc.) on other Gof tests like chi-square tests, we exploit the Kolmogorov-Smirnov (KS) test which is based on the empirical cumulative distribution function. The KS test computes the largest difference between the theoretical and the empirical distribution function [11]. Assuming that the random variable R_t represent the residual data and S_t represent the model distribution then the two sample KS test can be used to test whether the two underlying probability distributions differ. The KS statistics is given by:

$$M_t = max \ \ |R_t(x,y) - S_t\,(x,y)| \tag{5}$$

Where M_t is the least upper bound of all point wise differences $|R_t(x,y) - S_t\,(x,y)|$. If the sample comes from the same distribution then M_t converges to 0. The hypothesis regarding the residual data is rejected if the test statistics M_t is greater than the critical value.

The sonar data correspond to good quality (features appear very well), if the noise distribution in the data follows the Rayleigh distribution. If we assume that our residual measurements are governed by a particular distribution, we can make two working hypotheses about the distribution: The null hypothesis: H_0 : The observed distribution follows the Rayleigh distribution, while the alternative hypothesis: H_A : The observed distribution do not follow the Rayleigh distribution. The qualitative map must provide a quantitative value with confidence in the acceptance or rejection of the noise model. If the difference is less than a determined value then the agreement is satisfactory, but if the difference is much greater than the determined value then it is not satisfactory.

Based on the Kolmogorov-Smirnov test statistics given in equation (5), the null hypothesis regarding the observed distribution is rejected if the test statistics, (M_t value), is greater than the critical value (p-value) obtained from table [12]. The higher p-values represents the lower distance M_t value at the low range of the image, while the low p-values represent the higher distance M_t value at the far range of the residual image. Based on the p-value the sonar image is mapped into three different regions colored into Green, Blue and Red. The green and the blue colors are associated to Rayleigh distribution with higher p-values representing the very good and good regions while the red color represents the bad quality data with lower p-value.

3. EXPERIMENTAL RESULTS AND DISCUSSION

In this section we present the experimental results of the proposed approach. All the experimental tests are performed on sonar images acquired using DAURADE AUV robot. The sonar images are processed based on intensity data, where the darker gray color means low intensity of backscattering and the brighter color means high intensity of back scattering.

Fig. 1(a) is an example of raw sonar image acquired by DAURADE AUV, Fig. 1(b) represents the intensity correction of the raw sonar image using grey level normalisation process. In the Fig. 1(c), Fourier transform of the corrected sonar image is presented. The FFT image is used for the step of spectral filtering and residual extraction using Butterworth filter. Fig. 1(d) represents probability density function (PDF) of Fig. 1(c) for each 40 pixels of the range of the image. Fig. 2(a) is an example of filtered sonar image using Lee filter (standard deviation =9). Fig. 2(b) represents the PDF of Lee filtered image versus range image. In the Fig. 2(c), residual of Lee filtered image (residual=corrected sonar image – Lee filtered image) is presented. Fig. 2(d) represents the probability density function of the residual Lee filtered image versus range sonar. Fig. 2(e) represents the corrected sonar image filtered using Butterworth filter, Fig. 2(f) represents PDF of Butterworth filtered image versus range image. In the Fig. 2(g), the residual of Butterworth filtered image is shown and Fig. 2(h) the probability density function of the residual Butterworth filtered image versus range sonar is given. It can be observed that the

PDF in Fig. 2(h) is not constant and varies at far range of the sonar data, thus justifies the proposed approach by analysing the residual obtained through spectral filtering.

In order to show the quality mapping, we demonstrate the proposed approach on the famous sonar image of 'Swansea' acquired by DAURADE AUV given in Fig. 3. Fig. 4 represents a zoomed region of Fig. 3 given in rectangular red block. It can be observed from these two figures that the noise level varies from the near range to the far range of the sonar data. According to the proposed approach, the sonar image is partitioned into 64x64 size windows and transformed into frequency domain using Fourier transform. A high-pass Butterworth filter of order $n = 4$ with cutoff frequency r_0= 35 given by equation (1) is applied using equation (2). The residual image is obtained by taking the inverse Fourier transform.

Rayleigh distribution is fitted to the residual image and its parameters are estimated by using ML method. For the acceptance of the null hypothesis, two thresholds $\tau_1 = \frac{1.36}{\sqrt{64^2}.} = 0.02121$ and $\tau_2 = \frac{1.22}{\sqrt{64^2}.} = 0.019125$, are selected based on the significance level $(\alpha_1 = 0.05\ and\ \alpha_1 = 0.10)$ of the KS test for Gof [12].

Fig. 5(a) represent the cumulative distribution of the residual data along with CDF of the fitted Rayleigh distribution with p-value greater than τ_1, Fig. 5(b) represent the CDF of the fitted Rayleigh distribution with p-value greater than τ_2 but less than τ_1, respectively. From Fig. 5, it can be observed that the fitted Rayleigh from the residual data and the fitted Non-Rayleigh distributions are disjoint.

Based on the two sample KS tests statistics, if the critical value p is greater than τ_1, the null hypothesis is accepted (the data follows the Rayleigh distribution) with strong support $(\alpha_1 = 0.05)$, which represent the good quality data being represented in green color blocks in the Fig. 6 and Fig. 7. Similarly, if the critical value p is greater than τ_2 and less than τ_1, again the null hypothesis that the data follows the Rayleigh distribution is accepted but with relatively less support $(\alpha_2 = 0.10)$, which is represented by the blue regions in the Fig. 6 and Fig. 7. The regions which are not following the Rayleigh distribution represent those far regions where the signal response is very weak, represented by the red regions in the Fig. 6 and Fig. 7. In Fig. 8, we give an example of the portion of mapping along with the corresponding p values and color representation.

The results of the proposed mapping approach can be confirmed by the visual inspection of the raw sonar image and the mapped data. It has been observed that false positives (regions not following Rayleigh distribution) are also detected which limits the capabilities of the proposed approach in complex environments with high coefficient of variation.

4. CONCLUSION

Coherent imaging systems (like synthetic aperture radars, sonar, ultrasound and laser imaging) are commonly affected by multiplicative noise (also known as speckle noise). This article presents a novel statistical method for transforming the sonar data into quality mapping by exploiting the noise distribution in the sonar data. Based on the hypothesis that the noise follows the Rayleigh distribution, the proposed approach map the data into acceptable (good) region otherwise map the data into a non-acceptable (bad) regions. In the future, further analysis about the noise distributions in complex environments would make the proposed quality map more accurate and precise.

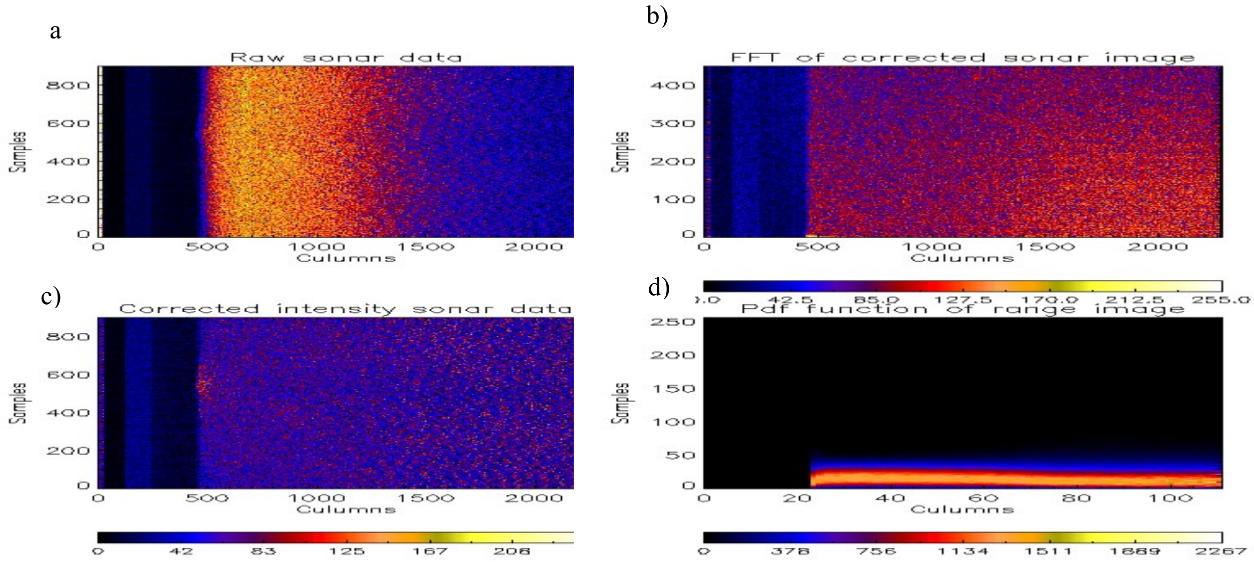

Fig. 1: a) Example of raw sonar image, b) Corrected of intensity of the raw sonar image, c) Fourier Transform of the corrected sonar image, d) Probability density function (PDF) vs. range sonar.

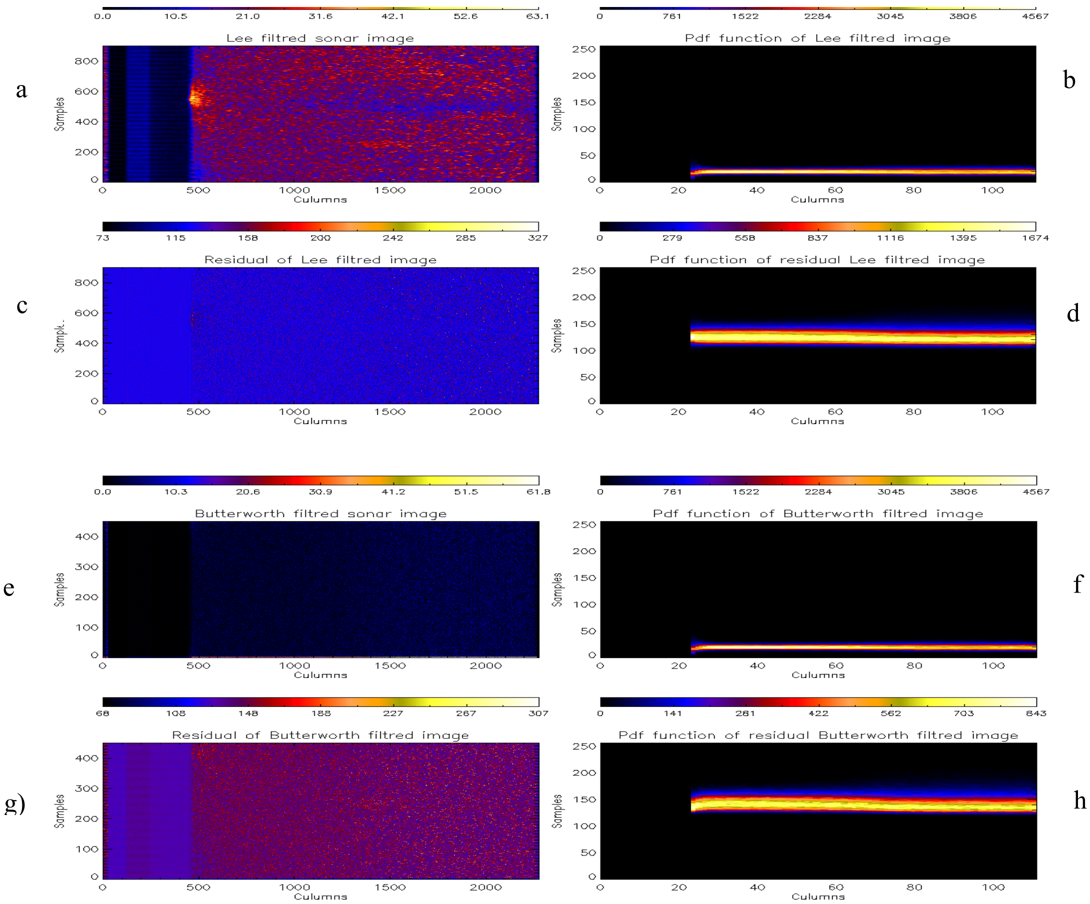

Fig. 2 : a) Corrected sonar image filtered using Lee filter, b) PDF of Lee filtered image vs. range image, c) Residual of Lee filtered image, d) Probability density function of the residual Lee filtered image vs. range sonar.
e) Corrected sonar image filtered using Butterworth filter, f) PDF of Butterworth filtered image vs range image, g) Residual of Butterworth filtered image, h) Probability density function of the residual Butterworth filtered image vs. range sonar.

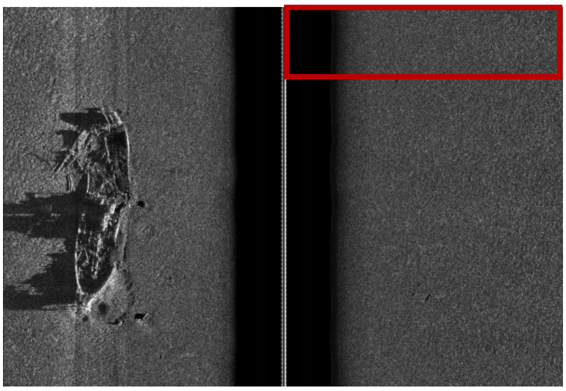

Fig. 3: Swansea Image

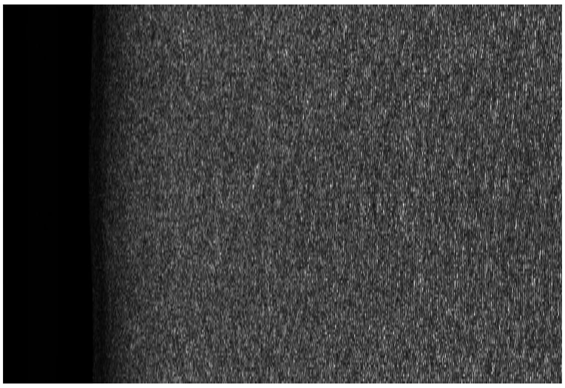

Fig. 4: Zoom region of the starboard in Fig. 1

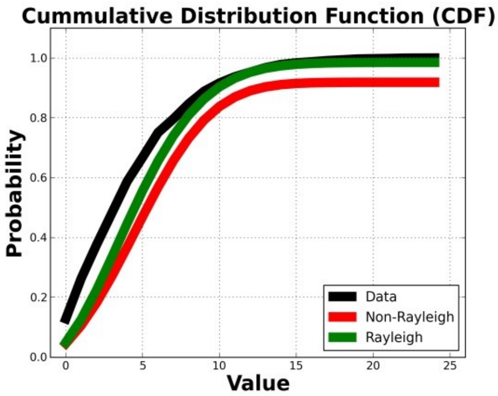

Fig. 5(a): CDF of the Residual data with Rayleigh ($p > \tau_1$) and Non-Rayleigh distribution

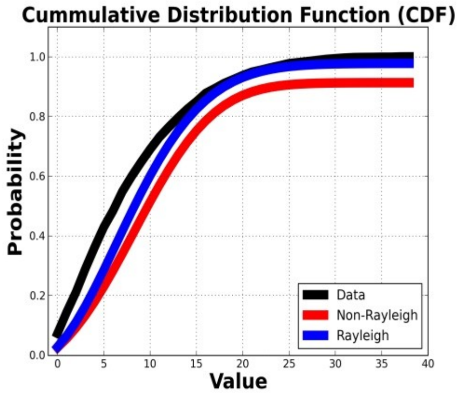

Fig. 5(b): CDF of the Residual data with Rayleigh ($\tau_2 < p < \tau_1$) and Non-Rayleigh distribution

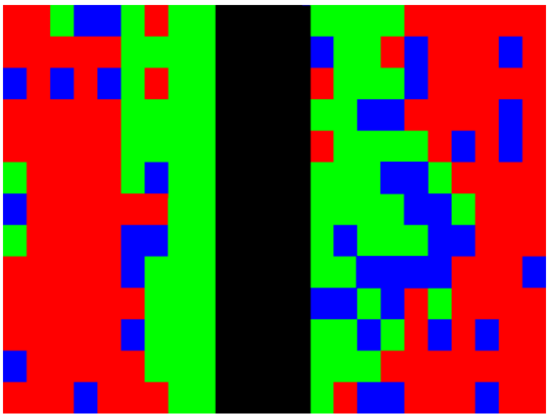

Fig. 6: Quality Mapping of the Swansea Image in Fig. 1

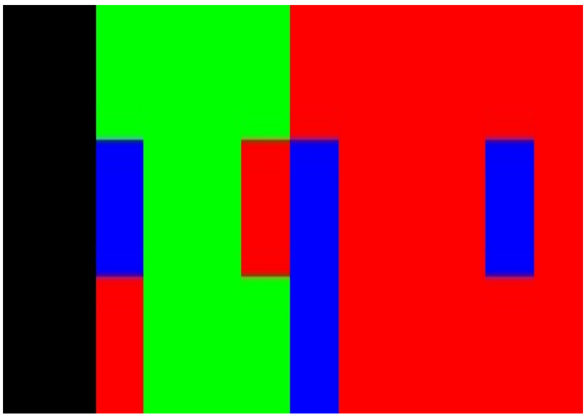

Fig. 7: Quality Mapping of the Image in Fig. 2

P	0	0	0.37	0.16	0.06	0.11	0.019	0.006	0.003	0.005	0.005	0.001
MAPPING												

Fig. 8: Quality mapping with the critical value p

REFERENCES

[1] B. Saevarsson, J. Sveinsson and J. Benediktsson, "Combined wavelet and curvelet denoising of SAR images," in IEEE International Conference on GEoscience and Remote Sensing Symposium, Anchorage, 2004.

[2] S. Ali and C. T. John, "Edge-detection-based filter for SAR speckle noise reduction," International Journal of Remote Sensing, vol. 33, no. 7, pp. 2296-2320, 2012.

[3] L. Gagnon, A. Jouan, "Speckle filtering of SAR Images- A comparative study between complex-wavelet-based and standard filters.

[4] K. Gupta and R. Gupta, "Despeckle and geographical feature extraction in SAR images by wavelet transform", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 62, no. 6, pp. 473-484, 2007.

[5] J. Maycock, B. M. Hennelly, J. B. Mc Donald, T. J. Naughton, Y. Frauel, A. Castro and B. Javidi, "Speckle reduction using the discrete Fourier filtering technique," in Machine Vision and Image Processing Conference, Ireland, 2007.

[6] David P. Williams, "AUV-enabled adaptive underwater surveying for optimal data collection", Intelligent Service Robotics, Springer-Verlag, vol. 5, no. 1, pp. 33-54, 2012.

[7] J. Bioucas-Dias and M. A. T. Figueiredo, "Multiplicative Noise Removal Using Variable Splitting and Constrained Optimization," IEEE Transactions on Image Processing, vol. 19, no. 7, pp. 1720-1730, 2010.

[8] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Prentice Hall, 3rd Edition, USA, 2007.

[9] A. Papoulis, S. Pillai, Probability, Random Variables and Stochastic Processes, McGraw-Hill, 4th Ed, USA, 2001.

[10] B. C. Gupta, I. Guttman, Statistics and Probability with Applications for Engineers and Scientists, John Wiley & Sons, USA, 2013.

[11] T. B. Arnold, J. W. Emerson, "Nonparametric Goodness-of-fit Tests for Discrete Null Distributions", The R Journal, vol. 3, no. 2, pp. 34-39, 2011.

[12] Kolmogorov-Smirnov Test [online], Available http://www.cas.usf.edu/~cconnor/colima/Kolmogorov_Smirnov.htm, (Accessed, 20 March 2014).