Rényi Entropy Filter for Anomaly Detection with Eddy Current Remote Field Sensors

Farid Sheikhi, Davide Spinello, *Member, IEEE,* Wail Gueaieb, *Senior Member, IEEE,*

***Abstract*—We consider a multi-channel remote field eddy current sensor apparatus, which is installed on a mobile robot which is deployed in pipelines with the mission of detection defects. Features in raw sensory data that are associated with defects could be masked by noise and therefore difficult to identify in some instances. In order to enhance these features that potentially identify defects, we propose an entropy filter that maps raw sensory data points into a local entropy measure. In the entropy space, data is then classified by means of a thresholding procedure based on the Neymnan-Pearson criterion. The effectiveness of the algorithm is demonstrated by applying it to different data sets obtained from field trials.**

***Index Terms*—Robot sensing systems; Signal mapping; Filtering algorithms; Entropy.**

I. Introduction

Water and energy sources are playing an important role in the world economy and are in an ever increasing demand. Their transportation mainly relies on a vast network of pipelines extending over millions of miles worldwide [1]. This makes the process of maintaining and servicing these pipelines in an autonomous and timely fashion, not only a very challenging task, but also an utmost necessity to avoid environmental disasters in the case of leakage, for instance.

Technology in this field is developing rapidly in order to provide efficient and quick inspection of pipelines [2], [3], [4], [5], [6], [7], [8]. The majority of pipeline inspection gauge (PIG) technology developers rely on nondestructive testing (NDT) to quickly and economically detect and characterize degradation and damage. There are many challenges that inspection tool developers must overcome. They must build inspection tools that survive the pipeline environment, meet specifications for long distance pipelines with high operational speeds, deal with pipeline abnormalities such as tight bends and obstruction in the pipe, and keep the measurement sensors in a good operating condition [9].

Many researchers attempted to tackle these challenges using different types of sensors and testing technologies. Due to profusion of metallic gas and oil pipes, magnetic sensors such as magnetic flux leakage, magnetic particle, and eddy current sensors [10], [11], are often regarded as most suitable for inspecting such pipelines. They are capable of detecting both internal and external defects. Their functioning principle is simple. Magnetic PIGs magnetize the pipe as they travel along it. A magnetic field related signal is captured by a group of transducers that are uniformly distributed around the inside circumference of the pipe. A difference in the levels of magnetic field penetration can be correlated to the existence of local flaws or defects [12]. The flaw can be due to corrosion, weld defects, cracks, fatigue, or deformation [13], [14], [15]. The resultant data collected from this process is typically characterized by low signal to noise ratio. Therefore, relevant information associated with the features to be identified is often masked by noise. This paper contributes to the effort of an ongoing research on developing effective techniques on extracting and analyzing relevant anomaly information from these excessively noisy signals.

Several types of sensors and pipe inspection technologies have been proposed in the literature. These include, for instance, visual testing, penetrant testing, magnetic particle testing, magnetic flux leakage, Hall-effect, radiographic testing, ultrasonic sensors, eddy current, thermal infrared, magnetic Barkhausen effect, and acoustic emission. More comprehensive reviews of these techniques are found in [16], [17], [18], [19], [20], [21] and in the references therein.

Acoustic and ultrasonic sensors were tested in [22], [23], [24]. A rotating optical geometry and infrared sensors were adopted in [25] and [26], respectively. Magnetic sensors, such as magnetic flux leakage [10], eddy current [11], [27], magnetic particle, and Hall-effect sensors, have been widely adopted in commercial gas pipe inspection and detection, taking advantage of the abundance of metallic utility pipes. This type of sensors are capable of detecting both internal and external defects, and are therefore suitable for nondestructive inspection methods. A comparative study between magnetic- and ultrasonic-based inspection techniques is reported in [28]. The authors in [29] proposed and compared three machine learning approaches, namely, support vector machine, kernelized principal component analysis, and kernelized partial least squares, to filter the sensor noise.

In this paper, we investigate the problem of extracting features from data generated by a multi-channel remote field eddy current sensor mounted on a snake-like mobile robot, Explorer-II. A full description of Explorer-II is provided in [30]. The problem is formulated in the framework of probabilistic classification theory, where data points have to be classified as noise or feature (anomaly). Some researchers used Shannon entropy [31] to filter noisy signals in order to distinguish features from background noise [32], [33], [34]. We extend this idea to the data series generated from a remote field eddy current sensor by considering the Rényi entropy [35] which is a one-parameter generalization of the

F. Sheikhi and D. Spinello are with the Department of Mechanical Engineering, University of Ottawa, Ottawa, ON, K1N 6N5 Canada. e-mail: {fshei046,dspinell}@uottawa.ca

W. Gueaieb is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON, K1N 6N5 Canada. e-mail: wgueaieb@uottawa.ca

Shannon entropy. The characterization of sensor noise and anomalies is done by thresholding the Rényi filter within the Neyman-Pearson decision making framework [36]. As a concept, the Rényi entropy has been used in many scientific applications, such as the target tracking of mobile networks and the reduction of data scope on ground-penetrating radar images for feature detection, for instance [37], [38].

The rest of the paper is organized as follows. In Section II we briefly introduce the concept and formalism of Rényi entropy, which delineate the theoretical framework of this paper. The Neyman-Pearson criterion to determine thresholds in hypothesis testing is discussed in Section III. In Section IV we present raw eddy current sensory data series obtained from different trials, that will be used to build and test the entropy filter. The entropy filter is presented in Section V, and it is applied in Section VI by showing how it operates and how it performs. Section VII is left for summary and conclusions.

II. Rényi Entropy

In information theory a typical measure of uncertainty associated to a random variable is the Shannon entropy, which is a way to express the expected value of the amount of information contained in a statistical population [39]. A generalized version of entropy has been proposed by Rényi [35], by introducing a one parameter family of entropies that includes Shannon's measure as a limit case. Formally, let X be a discrete random variable mapping to a finite dimensional events space $\{x_1, x_2, x_3, ..., x_N\}$ of cardinality N. Let $P(x_i)$ be a probability mass function over the set X, where $\sum_{i=1}^{N} P(x_i) = 1$. Rényi's entropy of order α ($\alpha > 0$, $\alpha \neq 1$) is defined by [35]

$$H_\alpha(X) = \frac{1}{1-\alpha} \ln \sum_{i=1}^{N} P(x_i)^\alpha = \frac{1}{1-\alpha} \ln(\mathbb{E}[P(X)^{\alpha-1}]) \tag{1}$$

By using L'Hôpital's rule it can be shown that H_α converges to Shannon entropy at the limit $\lim_{\alpha\to 1} H_\alpha$, that is

$$\lim_{\alpha\to 1} H_\alpha(X) = -\sum_{i=1}^{N} P(x_i) \ln P(x_i)$$

where the expression on the right-hand side conforms to the usual definition of Shannon entropy. Rényi entropy is a non-increasing function of α for all discrete random variables X, as it can be shown through differentiation with respect to α. The same property also intuitively follows from definition (1), as the limit for $\alpha \to 0$ is the case with highest entropy since $P(x_i)^0 = 1$ for all i.

III. Neyman-Pearson Criterion

Given a random variable H denoting an entropy data set, the goal is to achieve a systematic optimal method to decide which hypothesis generated that data. Consider the problem of identifying a signal masked by noise with the simple binary hypothesis test

$\mathcal{H}_0 : H$ corresponds to the entropy of sensor noise

$\mathcal{H}_1 : H$ corresponds to the entropy of known anomalies

Let parameters Θ_0 and Θ_1 be representatives of features in the respective hypothesis spaces. We assume that Θ_0 and Θ_1 form a disjoint class of the parameters space Θ. We denote the continuous distributions associated with the two hypotheses as F_0 and F_1 respectively,

$$\mathcal{H}_0 : H \sim F_0 = P(H|\Theta_0) \tag{2a}$$

$$\mathcal{H}_1 : H \sim F_1 = P(H|\Theta_1) \tag{2b}$$

where $P(H|\Theta_0)$ and $P(H|\Theta_1)$ denote class-conditional probabilities, defined by

$$P(H|\Theta_0) = \int_{\Theta_0} f_0(H)\, dH \tag{3a}$$

$$P(H|\Theta_1) = \int_{\Theta_1} f_1(H)\, dH \tag{3b}$$

and $f_0(H)$ and $f_1(H)$ are the probability density functions corresponding to noise and anomalies, respectively. Since the sensory data is inherently noisy the parameter chosen to identify anomalies is the entropy associated with data points. Therefore Θ_0 and Θ_1 identify, respectively, the entropy associated with noise and the entropy associated with anomaly. The probability of detection and the probability of false detection are, respectively,

$$P_D = 1 - \int_{\Theta_0} f_1(H)\, dH = \int_{\Theta_1} f_1(H)\, dH \tag{4}$$

$$P_F = \int_{\Theta_1} f_0(H)\, dH \tag{5}$$

The hypothesis is tested by evaluating the following ratio

$$\Lambda(H) \equiv \frac{f_1(H)}{f_0(H)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} \equiv \eta \tag{6}$$

where $\Lambda(H)$ is the likelihood ratio and η is a threshold. The threshold is assigned according to the Neyman-Pearson procedure [40] which does not require prior knowledge of hypothesis probabilities. Specifically, the value of the threshold is assigned by setting a constraint on the probability of false detection and it is formulated as a constrained maximization problem on the probability of detection:

$$\underset{\Theta_1}{max}\{P_D\}, \quad \text{such that } P_F \leq \epsilon \tag{7}$$

for some user-defined threshold ϵ which is usually set to 5%. The maximization is performed over the decision region Θ_1 and it selects the most powerful test subjected to the constraint on the probability of false detection [41]. The procedure for building the hypothesis test and the threshold is applied in Sec V-B.

IV. Data from Eddy the Current Sensor

In data series characterized by low signal to noise ratio, there are rare events that reveal specific features and are masked by noise. The entropy filter is based on the fact that such rare events carry more information than the background noise, and therefore the associated features in the data series can be extracted as they are enhanced by the action of the filter on the original data. In the present study, data series are generated by a remote field eddy current sensor [17] mounted

on a mobile robotic platform which performs non-destructive inspection of gas pipelines [30]. Sensor measurements are correlated to the presence of defects (anomalies) of the pipeline [42]. Multiple detectors distributed along the circumference of the pipe allow for a discretized coverage of the surface. As discussed in [43], the capability of detecting defects and the accuracy of detected entities depend on the quality of the eddy current signal, which is typically affected by different noise sources as tube support structures, corrosion products, changes in tube dimensions and geometry, and probe wobble and lift-off [42]. Additionally, non-uniform testing surface conditions and electronic noise from the test equipment can also be noise sources [43].

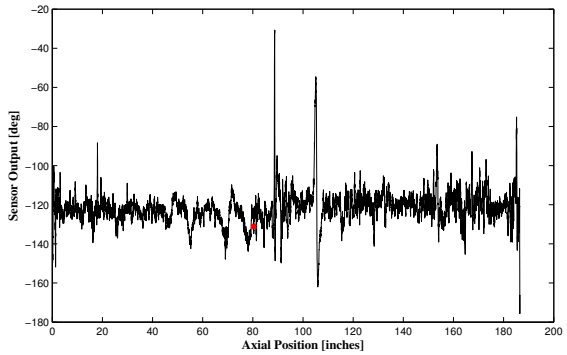

Fig. 1. Phase (in degrees) of the output of a single channel of the eddy current sensor.

A typical output of one channel of a remote field eddy current sensor is shown in Fig. 1. The vertical axis represents the phase (in degrees) of the sensor's output signal, while the horizontal axis represents a longitudinal coordinate (in inches) spanning a portion of a pipeline in which the sensor collected the data. The data series refers to a portion of a pipeline with few significant defects in the middle, with peaks at the boundaries most likely due to inaccurate contact of the sensor at the beginning and at the end of the operation. Raw sensory data from the remote field eddy current sensor is provided in the form of arrays of In-Phase (I) and Quadrature (Q) components, that are basically Cartesian coordinates of the phasor representation of the sensor output. In this work we consider the associated phase data, which is obtained by the relation $\phi = \mathrm{atan2}(Q, I)$ where atan2 is the arctangent function with two arguments [44], mapping to $[-\pi, \pi]$ by accounting for the quadrant in which each of the two arguments belongs. Phase data from an eddy current sensor has been used in [43] to study the detectability and sizing-accuracy of laboratory induced defects in nuclear steam generator pipes. Experimental evidence in [43] supports the link between enhanced detectability and increasing signal to noise ratio, and it is shown how in standardized laboratory conditions, pipes' surface polishing can strongly increase the signal to noise ratio. The scalar field ϕ is represented as a two dimensional array $\phi(i, j)$, where $i = 1, \ldots, N_\ell$ and $j = 1, \ldots, N_w$, respectively span the discrete axial position along the pipe and the circumferential position which is sampled by a bundle of data channels comprising the sensor (see the schematic in Fig. 2). Phase measurements are correlated to the thickness of the pipe and therefore to the presence of defects (anomalies) quantified by a change in thickness. According to the work in [43], the pipe surface finish (and therefore the contact between sensor and pipe inner surface) is a dominant noise factor. Therefore it is expected that in field conditions involving gas pipelines, naturally occurring corrosion and analogous non ideal surface conditions introduce a considerable amount of noise associated to non-ideal contact. The aim is to develop an algorithm that enhances low noise structures so that defects, that are characterized by higher information content than the background noise, can emerge and be identified.

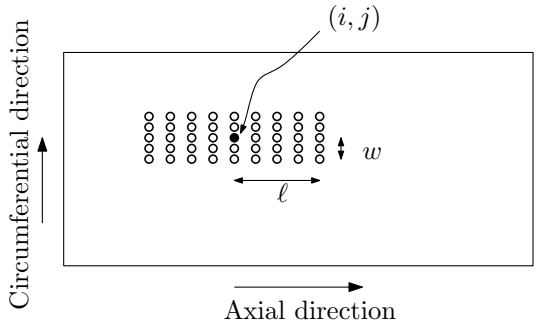

Fig. 2. Schematic of a rectangular window centered at the phase datum $\phi(i, j)$

V. Entropy Filter

The entropy filter presented in this section is structured in two parts: the first is a nonlinear integral transformation that maps the raw sensory data into an entropy measure; the second is a thresholding process that classifies the transformed data based on whether or not it represents an anomaly. In this work, we consider a binary classifier that is based on Neyman-Pearson criterion. The entropy map enhances the local information content of the data points, therefore better captures structured signals that were buried in noise in the original raw sensory data space.

A. Mapping to Local Entropy

In order to enhance relatively rare structures with higher information content than the background noise, raw sensory data series collected by the sensor are mapped into a local entropy space as follows. For each raw data point $\phi(i, j)$, $i = 1, \ldots, N_\ell$ and $j = 1, \ldots, N_w$, we define the subset (see the graphical representation in Fig. 2)

$$\Phi(i,j) := \{\phi(h,k) : h \in [\max(i-\ell, 1), \min(i+\ell, N_\ell)], \\ k \in [j-w \mod N_w, j+w \mod N_w]\} \tag{8}$$

that defines a window centered at ϕ_{ij} with sides $2\ell + 1$ and $2w+1$, for non-negative integers ℓ and w. The notation $(j-w) \mod N_w$ (and similarly $(j+w) \mod N_w$) refers to the usual modular arithmetic, introduced to account for the periodicity along the circumferential direction; in this case we offset the operation by 1, so that the reminder of the integer division $(j-w)/N_w$ maps to $[1, N_w]$ rather than to $[0, N_w - 1]$. The

cardinality of this data set is therefore $(2\ell + 1)(2w + 1)$. Intuitively, ℓ and w account for the interaction of a data point with neighbors: for $\ell = w = 0$, we have $\Phi \equiv \phi$ for all data points; whereas for $\ell \geq N_\ell$ and $w \geq N_w$, Φ includes the entire data set.

In order to compute the local entropy associated to a data point $\phi(i, j)$, the histogram distribution of the associated local data set $\Phi(i, j)$ is built with sample space given by the set $\Phi(i, j)$. The optimal number of bins is obtained by the optimization procedure discussed in Appendix A. To illustrate the process of optimal bin number determination, we select a data point from the set in Fig. 1 (red dot, located at the abscissa $81\,\text{in}$) and show in Fig. 3 the histogram with optimal number of bins corresponding to the minimum of the plotted cost function. The optimal number of bins depends on the cardinality of the sample space as well as the characteristics of data points, as detailed in Appendix A.

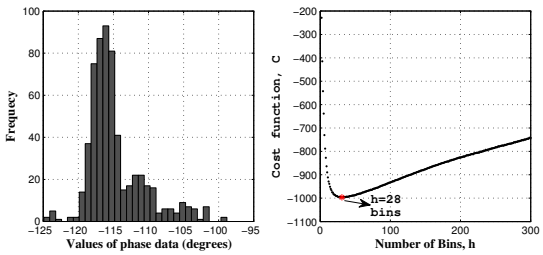

Fig. 3. Histogram of the sample space of one data point from Fig. 1 and related cost function to find the optimal number of bins

For the data point $\phi(i, j)$, let $k^\star$ be the optimal number of bins of size $h^\star$ associated with the histogram of the set $\Phi(i, j)$. Therefore, the sample space in the histogram representation is divided into $k^\star$ contiguous intervals delimited by the sequence $\min \Phi(i, j) + kh^\star$, for $k = 0, 1, \ldots, k^\star$, with corresponding mass probability

$$P(\bar{\phi}_k) \equiv$$
$$P(\bar{\phi}_k \in [\min \Phi(i, j) + kh^\star, \min \Phi(i, j) + (k + 1)h^\star]) \quad (9)$$

where $\bar{\phi}_k \in \Phi(i, j)$ is a representative of the k-th bin. Therefore the local Rényi associated to the data point ϕ_{ij} is the map defined by

$$H(\phi_{ij}) = \frac{1}{1 - \alpha} \ln \sum_{k=0}^{k^\star - 1} P(\bar{\phi}_k)^\alpha. \quad (10)$$

In what follows, the results are computed for windows defined by $\ell = 100$ and $w = 1$. Numerical experiments showed that these values yield a satisfactory compromise between preserving the local features of the data when mapped to the entropy space, and the correlation of neighboring data points. A discussion on the influence of these parameters is provided in Section VI-B.

B. Threshold for Data Classification

The anomaly detection with the entropy filter is completed with a threshold to discriminate the noise from infrequent structures with higher information content, as they are associated to features to be detected. The threshold is determined by using the Neyman-Pearson criterion which assigns the value corresponding to the maximum detection probability achievable for a given false detection rate [36], [45] (see Section III). As evident from equation (6) the classification test needs the specification of two probability density functions f_0 and f_1, that characterize respectively the noise and the anomalies (or features to be detected). Here, we consider Gaussian approximations which highly simplify the procedure to obtain the threshold, as detailed below. Gaussian distributions are a good approximation for data series characterized by noise, and therefore in the specific case for data collected in portions of pipelines with no or very few defects. The approximation needs to be refined for data sets with a considerable amount of defects, in which higher peaks and fat tails have to be captured. This refinement is the object of current work.

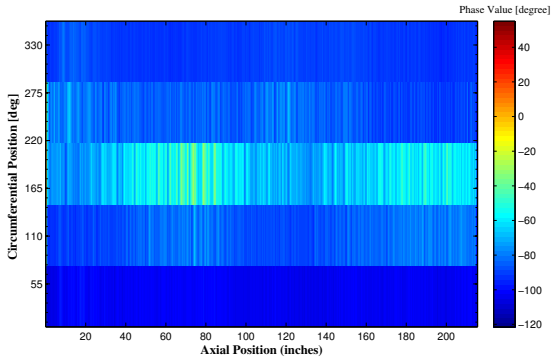

Fig. 4. Two-dimensional phase data set from raw sensory data acquired on a pipeline without anomalies

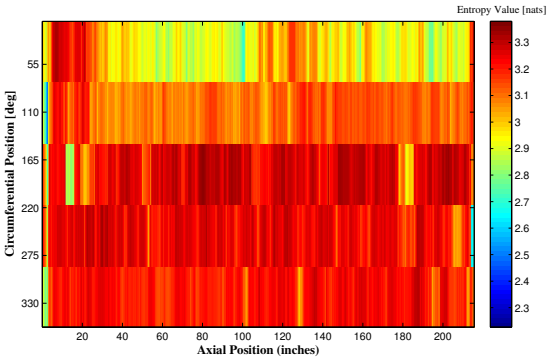

Fig. 5. Local Rényi entropy corresponding with the data in Fig. 4 computed with $\alpha = 0.5$, $\ell = 100$, and $w = 1$

To determine f_0, we consider a data set from a field trial on a portion of a pipeline that was knowingly anomaly free. The phase data is shown in Fig. 4. This data is in cylindrical coordinates and is measured in a lab run from a pipeline with no anomalies. One major source of noise is the non

ideal contact of the sensor with the pipe. The figure shows an almost uniform distribution of phase values over the entire space with few light strips that are possibly due to a local faulty condition of sensors. The phase data is mapped to local entropy according to the expression (10), where each data point is centered on a window defined by parameters $\ell = 100$ and $w = 1$ (see Section V-A). The density plot of the entropy data is given in Fig. 5. We define the sample space Φ_0 of cardinality N_0, that includes the local entropy data shown in Fig. 5, and compute the corresponding sample moments (sample mean and sample variance)

$$\mu_0 = \frac{1}{N_0}\sum_{i,j} H(\phi_{ij}^0), \quad \sigma_0^2 = \frac{1}{N_0}\sum_{i,j}\left(H(\phi_{ij}^0) - \mu_0\right)^2 \quad (11)$$

where $\phi_{ij}^0 \in \Phi_0$ is a data point in Fig. 4 belonging to the sample space Φ_0. The plot in Fig. 6 shows the discrete probability density function (dots) obtained from the normalized histogram of the data in Fig. 5. The solid line in the same plot is the normal probability density function $\mathcal{N}(\mu_0, \sigma_0)$, which is used as a continuous model of the discrete one. To determine the threshold and to classify the data mapped into the local entropy space we assume $f_0 \equiv \mathcal{N}(\mu_0, \sigma_0)$ where $\mu_0 = 3.152\,\text{nats}$ and $\sigma_0 = 0.081\,\text{nats}$.

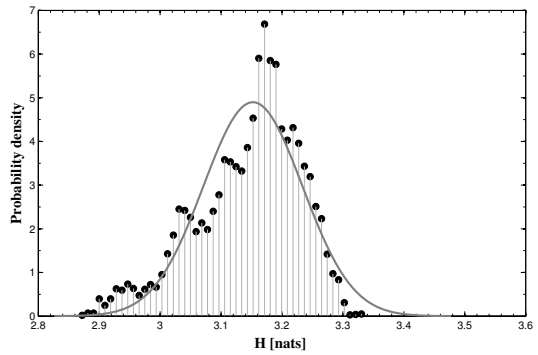

Fig. 6. Discrete (dots) and approximating Gaussian (solid line) probability density functions for the entropy data set associated with sensor noise.

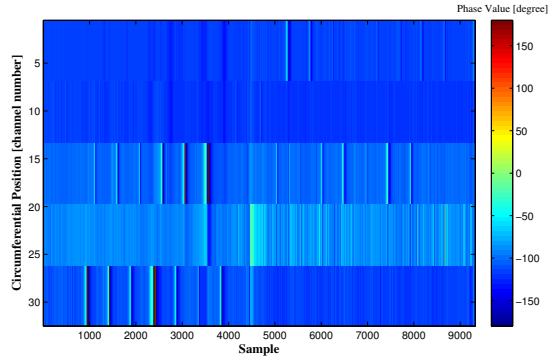

Fig. 7. Phase data set in cylindrical coordinates from a portion of pipeline with known, machined defects

The characterization of the density f_1 follows the same steps, except that in this case we consider a data set related

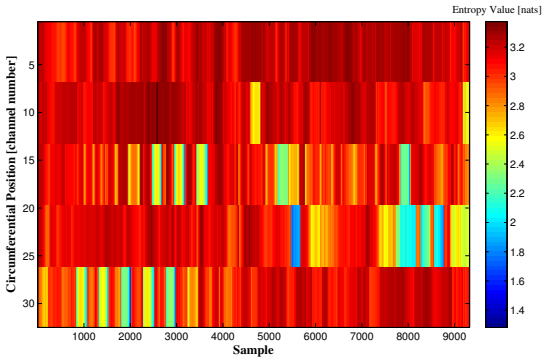

Fig. 8. Local Rényi entropy for the data in Fig. 7 computed with $\alpha = 0.5$, $\ell = 100$ and $w = 1$

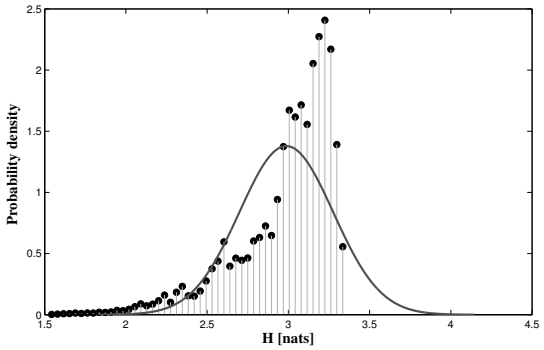

Fig. 9. Discrete (dots) and approximating Gaussian (solid line) probability density functions for the entropy data set associated with known defects.

to sensor measurements from a portion of pipeline containing anomalies that were introduced in a controlled way by machining pipe segments, see Fig. 7. Even on a portion of pipeline with artificially introduced damages, the number of sensor measurements associated to anomalies is much smaller than the number of sensor measurement associated to undamaged areas (noise). Therefore, by including all data points we would obtain statistical moments not significantly different than the ones computed for f_0. Hence, only parts with defects are selected in order to have a sample space that is representative of defects. Specifically, this data set spans eight different anomalies ranging from 0.5 to 3 inch in length and width, and 5% to 30% of pipe wall thickness in depth. The density plot of the local entropy computed with $\ell = 100$ and $w = 1$ is given in Fig. 8. The discrete distribution from the normalized histogram of the anomaly data in Fig. 8 and the approximating Gaussian density are plotted in Fig. 9. The sample statistical parameters of the Gaussian density function are calculated as

$$\mu_1 = \frac{1}{N_1}\sum_{i,j} H(\phi_{ij}^1)\ , \quad \sigma_1^2 = \frac{1}{N_1}\sum_{i,j}\left(H(\phi_{ij}^1) - \mu_1\right)^2 \quad (12)$$

where ϕ_{ij}^1 is a data point in the sample space Φ_1 (with cardinality N_1) comprised of local entropy values extracted to represent anomalies. The computed values of the parameters are $\mu_1 = 2.987\,\text{nats}$ and $\sigma_1 = 0.289\,\text{nats}$. In this case, since the data is not purely noisy, the probability density is not

TABLE I
PARAMETERS FOR THE PROBABILITY DENSITY FUNCTIONS f_0 AND f_1 FOR DIFFERENT VALUES OF α, COMPUTED FOR $\ell = 100$ AND $w = 1$

α	μ_0	σ_0	μ_1	σ_1
0.1	3.321	0.0577	3.210	0.2430
0.5	3.152	0.081	2.987	0.289
1	2.970	0.1179	2.774	0.3602

TABLE II
THRESHOLDS H^- FOR DIFFERENT VALUES OF α, COMPUTED FOR $\ell = 100$ AND $w = 1$

α	H^-
0.1	3.213
0.5	3.004
1	2.758

obviously Gaussian; however, for simplicity of treatment we consider f_1 to be approximated by a normal distribution with parameters μ_1 and σ_1, see Fig. 9. The statistical parameters in f_0 and f_1 are summarized in Table I. As can be noted, the mean of distributions f_0 and f_1 are relatively close.

Given the Gaussian densities f_0 and f_1, the likelihood ratio for hypothesis testing can be extracted from (6) and is explicitly written as

$$\Lambda(H) = \exp\left(-\frac{1}{2}\left(\frac{H-\mu_1}{\sigma_1}\right)^2 + \frac{1}{2}\left(\frac{H-\mu_0}{\sigma_0}\right)^2\right) > \eta\frac{\sigma_1}{\sigma_0} \tag{13}$$

Taking the logarithm of both sides, we obtain a quadratic inequality in H that defines the detection region (critical region) as a function of η

$$\left(\frac{H-\mu_0}{\sigma_0}\right)^2 - \left(\frac{H-\mu_1}{\sigma_1}\right)^2 - 2\ln\left(\eta\frac{\sigma_1}{\sigma_0}\right) > 0 \tag{14}$$

Let

$$a = \frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} \tag{15a}$$

$$b = -2\left(\frac{\mu_0}{\sigma_0^2} - \frac{\mu_1}{\sigma_1^2}\right) \tag{15b}$$

$$c = \frac{\mu_0^2}{\sigma_0^2} - \frac{\mu_1^2}{\sigma_1^2} - 2\ln\left(\eta\frac{\sigma_1}{\sigma_0}\right) \tag{15c}$$

The left-hand side of (13) can now be rewritten as $aH^2 + bH + c$. Let $H^-(\eta)$ and $H^+(\eta)$ be the roots of this quadratic function of H, with $H^- \leq H^+$. The detector in (13) therefore dictates the following detection region

$$\Theta_1(\eta) = \begin{cases} H < H^-(\eta) \cup H > H^+(\eta) & \text{if } a > 0 \\ H^-(\eta) < H < H^+(\eta) & \text{if } a < 0 \end{cases} \tag{16}$$

The value of the threshold η is found by numerically solving (5) with the bisection method having the upper bound of P_F (i.e. ϵ in (7)) set to 5%, which is a value often adopted in the existing literature. For a false alarm rate P_F upper-bounded by 5%, the values of the threshold are given in Table II for different Rényi entropy measures with parameters $\ell = 100$ and $w = 1$. The parameter H^+ is not considered here since anomalies cannot have a lower entropy than noise. The entropy filter therefore reduces to the two-phase algorithm comprised of the local entropy computation followed by the classification based on threshold H^-.

C. Summary of the Entropy Filter

The entropy filter presented in this section is structured into two main procedures:

1) Mapping the row sensory data into a suitable entropy space (Rényi entropy).
2) Applying a threshold that classifies the data in the entropy space.

The mapping into the Rényi entropy space is executed through the following steps:

1a) Acquire row sensory data.
1b) Associate a neighbor to each raw sensory data point according to Eq. (8).
1c) Build the local histogram associated to the neighbor of each data point, which determines mass probabilities given by Eq. (9). A procedure to obtain the optimal number of bins for the histograms is presented in the Appendix.
1d) Apply the map (10) that associates each raw sensory data point to the corresponding local entropy.

The data classification procedure can be executed in different ways, depending on the nature of the problem and on the theoretical tools adopted. In this work we adopt a simple classifier based on Neyman-Pearson criterion, which allows to calculate a threshold by maximizing the probability of detection for a given probability of false detection, in this case set to 5%. The steps to obtain the threshold can be summarized as

2a) Identify features of the classifier that define the partitioning of the data. For example, in the case studied here, data points have to be classified to identify low entropy features in noisy data sets, or otherwise anomaly versus noise.
2b) Identify data sets that characterize each partition.
2c) Map the characterizing data sets into local entropy space as explained above.
2d) Calculate the threshold according to the procedure delineated in Section V-B, which is based on Neyman-Pearson criterion with constrained probability of false detection set to 5%.
2e) Classify the data through the likelihood ratio test in eq. (13) or the log-likelihood ratio test in eq. (14).

VI. RESULTS AND DISCUSSION

A. Anomaly Detection

In order to test the consistency of the entropy filter algorithm, we apply it to the same data sets used to obtain the statistical parameters for the densities f_0 and f_1, see Section V-B. This process should, after thresholding the local entropy data, lead to capturing almost all noise when applied to the data set used to compute f_0, and to detecting the majority of the anomalies when applied to the data set used to obtain

f_1. Figs. 10 and 11, respectively, illustrate the raw phase of the signal fed by channel 2 of the eddy current sensor applied to the purely noisy data set used to build f_0, and the corresponding local entropy computed with $\alpha = 0.5$, $\ell = 100$, and $w = 1$. The threshold is indicated in Fig. 11 with a horizontal dashed line. The entropy filter correctly classifies all data as noise and no defect or anomaly is detected, since the threshold is lower than the local entropy throughout the pipe segment. The classification with the threshold H^- in Table II, obtained through the Neyman-Pearson criterion, gives the result in Fig. 12, where the output of the entropy filter is represented by the density plot which is obtained by linearly interpolating data from the contiguous channels to reconstruct a two dimensional profile whose support is the surface defining the portion of pipeline. In this case, 83.4% of the entropy mapped data is higher than the threshold, and is therefore classified as noise.

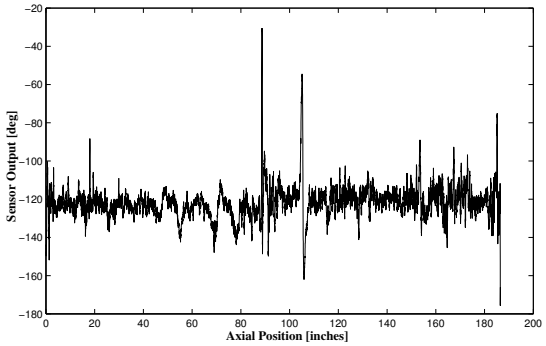

Fig. 10. Raw phase from channel 2 of the eddy current sensor.

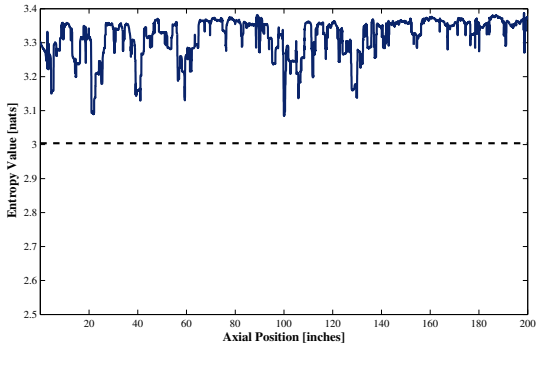

Fig. 11. Entropy of a single sensor channel referring to a noisy data set, and the threshold value used with the entropy filter.

By applying the entropy filter with the same values of α and w to the data set used to obtain the density f_1 we obtain the result in Fig. 13, where the output is restricted to a region of a pipeline containing known defects with centroids identified by the dots. Critical regions associated to defects are represented by red and yellow shades and are clearly detected by the filter. The relation between the threshold and the entropy data from channel 24 of the same data set in Fig. 13 is shown by the plot in Fig. 14; in this case, several regions around defects are clearly below the threshold and therefore identified as defects.

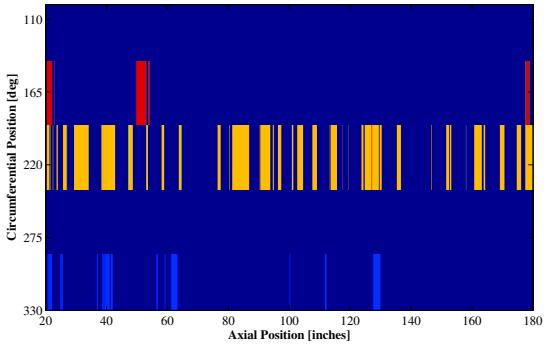

Fig. 12. Application of the entropy filter to a multichannel data set from a portion of a pipeline without defects. Blue regions refer to noise, whereas warm colored regions are false critical regions.

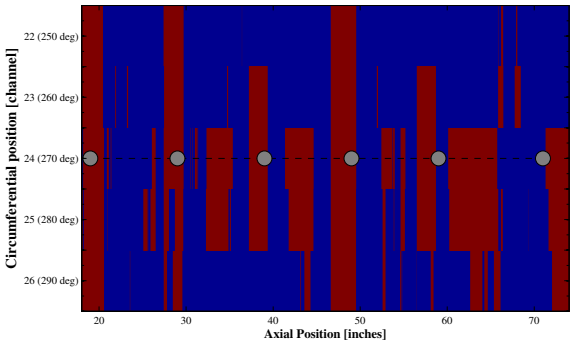

Fig. 13. Application of the entropy filter to a multichannel data set from a portion of a pipeline with known anomalies (gray circles). Red and yellow regions are critical with respect to the entropy filter, and therefore identify defects against the noise represented in blue.

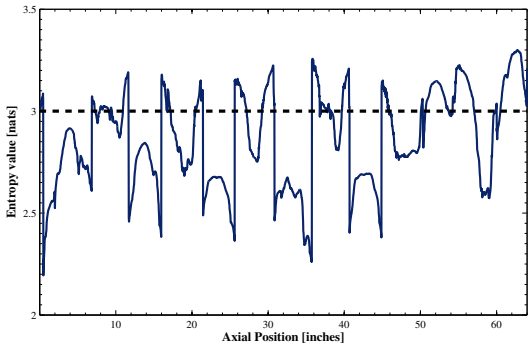

Fig. 14. Entropy data (continuous line) from channel 24 across a region with known anomalies, and threshold value used with the entropy filter (dashed line).

By applying the entropy filter with $\alpha = 0.5$, $w = 1$ and the threshold H^- in Table II to a data set acquired on a portion of a pipeline with welded joints, we obtain the result in Fig. 15, where the welding structure along the circumferential direction is well identified by the red region. The axial pattern shown in the figure may be due to faulty contact between a sensory channel and the surface of the pipeline.

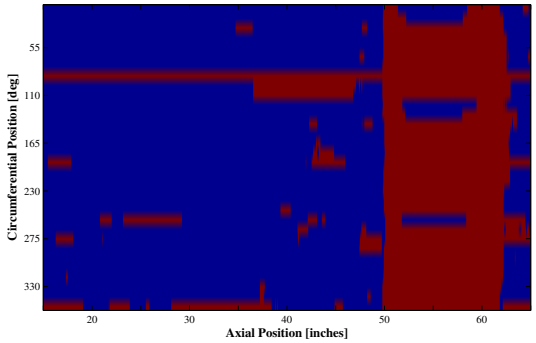

Fig. 15. Application of the entropy filter to a multichannel data set from pipeline segment with welded joints. Red regions are critical with respect to the entropy filter, and therefore identify defects (in this case welded joints).

TABLE III
PROBABILITY OF DETECTION P_D COMPUTED FOR DIFFERENT PARAMETERS AND FIXED FALSE ALARM RATE $P_F = 5\%$

	$\ell = 50$	$\ell = 100$	$\ell = 150$
$\alpha = 0.1$	32.8%	67.6%	83.7%
$\alpha = 0.5$	32.5%	64.3%	79.6%
$\alpha = 1.0$	31.4%	58.6%	74.9%

B. Influence of Parameters

By referring to eq. (1) we numerically investigate the effect of the Rényi parameter α and of the window size parameter ℓ for constant probability of false alarm $P_F = 5\%$ and constant number of contiguous sensor channels included in the computation of the local entropy ($w = 1$). The parameter ℓ is varied by taking the discrete values $\{50, 100, 150\}$ while α is varied by taking the discrete values $\{0.1, 0.5, 1.0\}$.

Table III summarizes the probability of detection P_D corresponding to different combinations of α and ℓ. The probability of detection increases as the window size ℓ increases, since a larger window size corresponds to more data and therefore more information in the neighborhood of a given raw data point. However, it should be noted that a larger window size is computationally more expensive. By choosing the window size too large the entropy filter looses resolution and accuracy as it cannot discriminate between data points (at the limit, the entropy would be uniform as the window size would include the entire data set). On the other hand, smaller values of α correspond to larger probability of detection; indeed as α approaches zero, the Rényi entropy increasingly weighs all possible events more equally, regardless of their probabilities, see eq. (1). This is a similar situation to a large window size, as in both cases the system tends to equal mass probability for data points, and therefore more uniform and higher entropy. The combination of parameters $\alpha = 0.5$ and $\ell = 100$ used in the previous Section gives $P_D = 64.3\%$.

Density plots showing the output of the entropy filter applied to the same data set used to generate the plot in Fig. 13 are shown in Fig. 16 for different combinations of α and ℓ. Plots illustrate the effects of parameters discussed above, in particular showing the loss of resolution when the window size for the computation of the local entropy increases by increasing significantly the parameter ℓ. Note that, as indicated in the figure caption, the numerical value of the threshold H^- is different for each combination, as the procedure explained in Section V-B is repeated for each set of entropy parameters.

VII. SUMMARY AND CONCLUSION

We proposed and tested a data processing and filtering algorithm to detect relevant features, often in the form of events with low probability, from data series in which such features may be masked by noise. Row data is acquired by a remote field eddy current sensor on board a mobile robot for non-destructive inspection of gas pipelines, and the eventual presence of defects can be correlated to the signal characteristics (magnitude and phase) of the data. In order to identify rare events masked by noise, we take advantage of the fact that such events are more ordered, in an information theoretical sense, than the background noise intrinsically associated to the sensor. Therefore we map the data to the entropy space by means of an entropy filter. The entropy filter computes the entropy associated to every data point, accounting for a user defined neighborhood or subset of the original data. Once the original data is mapped into the entropy space, the filter is completed by a classification phase that discriminates between noise and anomalies (features), through a thresholding procedure based on Neyman-Pearson criterion.

The operation of the filter is illustrated by testing it on different data sets. As expected, the anomalies introduced in a controlled setting are correctly detected. In addition, the correlation between phase data and entropy data shows that high phase peaks appear in regions that the filter classifies as anomaly, and it is important test for uniformity of the method. Moreover, the anomaly represented by a weld on a different pipeline is sharply detected, and a pipeline with no known anomalies is predicted to be anomaly free, except the wrong detection associated to one specific channel of the sensor that may be attributed to faulty local conditions of the sensor.

The study of the influence of characteristic parameters of the filter shows the trade-off effect of the neighborhood size on the accuracy of the filter. The observations indicate that too large neighborhoods result into lack of discrimination of data points in terms of information content and into computationally expensive post processing processes. On the other hand, too narrow neighborhoods result in the influence of local fluctuations in the output of the filter. On the other hand, deriving the correlation between size of actual anomalies and the detected ones is beyond the scope of this works and asymmetric study in this direction is left for future work. Overall, the proposed algorithm proves to be effective in introducing a layer of automation in the processing of large sets of experimental data, to help identifying critical regions that eventually may be inspected by experts, or analyzed in a refined way.

APPENDIX A
DETERMINATION OF THE OPTIMAL HISTOGRAM BIN SIZE

A histogram is one of the simplest methods for non-parametric density estimation. It is very efficient when the

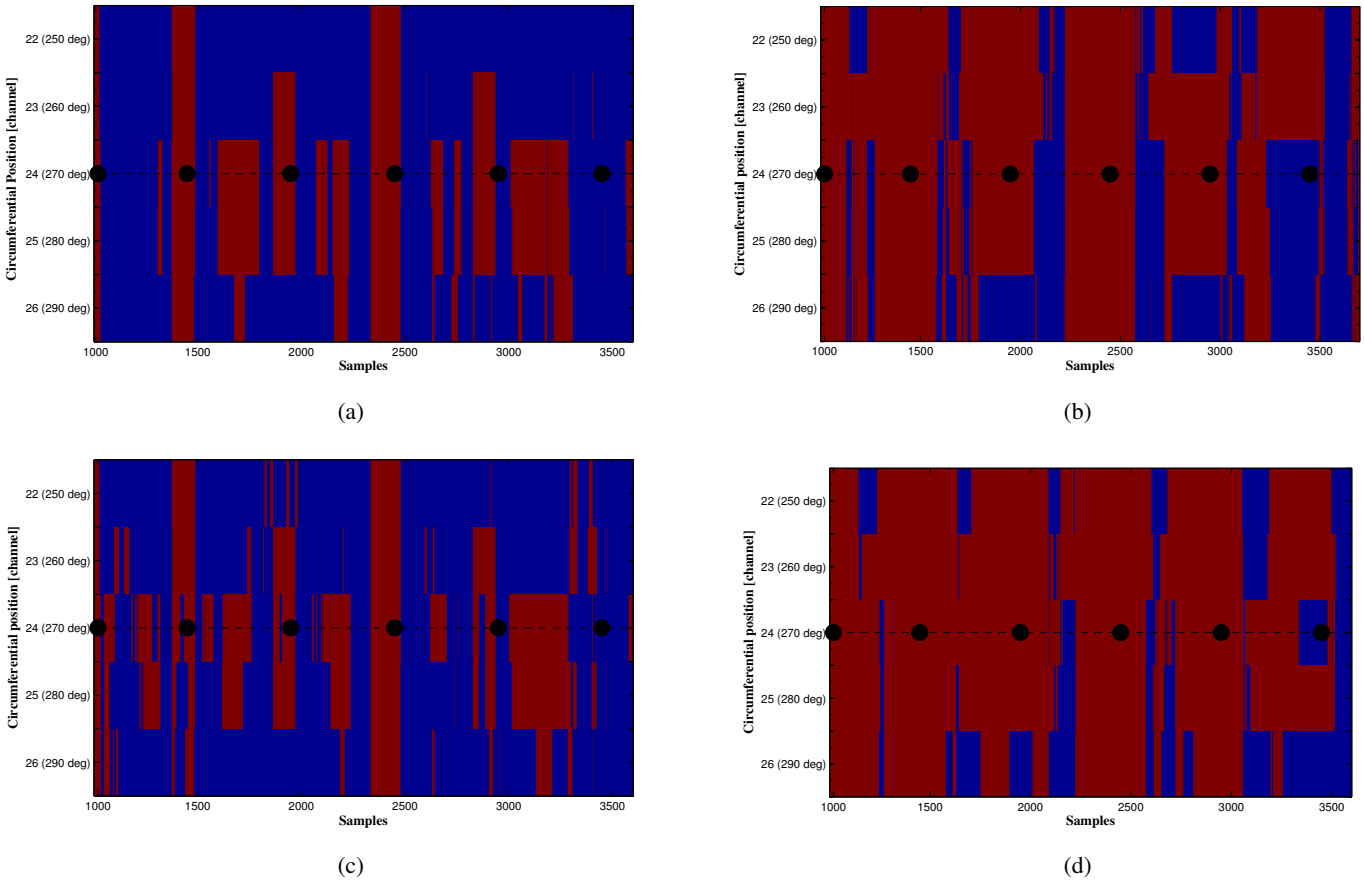

Fig. 16. Application of the entropy filter with (a) $\alpha = 0.1$, $\ell = 50$, and threshold $H^- = 3.015$; (b) $\alpha = 0.1$, $\ell = 150$, and threshold $H^- = 3.846$; (c) $\alpha = 1$, $\ell = 50$, and threshold $H^- = 2.592$; (d) $\alpha = 1$, $\ell = 150$, and threshold $H^- = 3.375$.

data sample is one- or two-dimensional [41]. A critical aspect in building the histogram of a data set is the choice of the number of bins, as this parameter controls the trade-off between the resolution in sample space and frequency fluctuations associated to data self correlations and [46].

In this work we implement an optimization method, which is based on the minimization of the expected L^2 loss between the histogram and the underlying density function [46]. The assumption is that data points are sampled independently (Poisson process). In most of classical density estimations, the sampling size is fixed. However, under Poisson assumption, the total data size and the number of events occurring in a fixed interval are not fixed but rather obey Poisson distribution. Therefore, it is more adaptable than the histograms constructed on samples with a fixed number [46]. For a one-dimensional data set, the algorithm is structured as follows:

1) Divide the data set X into k bins of width h, and build the histogram by counting the number of events m_i in the i-th bin.
2) Calculate the sample mean and sample variance of the number of events as follows:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} m_i \qquad \sigma = \frac{1}{N}\sum_{i=1}^{N}(m_i - \mu)^2$$

3) Find the solution $h^\star$ of the following unconstrained optimization problem

$$h^\star = \underset{h}{\operatorname{argmin}}\, C(h)$$

with cost $C(h)$ given by

$$C(h) = \frac{2\mu - \sigma^2}{h^2}$$

4) Calculate the optimal number of bins $k^\star$ as

$$k^\star = \left\lceil \frac{\max(X) - \min(X)}{h^\star} \right\rceil$$

where $\lceil \cdot \rceil$ is the ceiling function.

ACKNOWLEDGMENT

This research has been partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), grant number EGP 402129-10, and the Federal Economic Development Agency for Southern Ontario (FedDev Ontario).

The authors would also like to thank InvoDane Engineering Ltd. for their support to the project.

REFERENCES

[1] R. Tubb, "2011 worldwide pipeline construction report," *Pipeline and Gas Journal*, vol. 238, no. 1, pp. 18–38, 2011.

[2] L. Udpa, S. Mandayam, S. Udpa, Y. Sun, and W. Lord, "Developments in gas pipeline inspection technology," *Materials Evaluation*, vol. 54, no. 4, 1996.

[3] J. Nestleroth and R. J. Davis, "Application of eddy currents induced by permanent magnets for pipeline inspection," *NDT and E International*, vol. 40, no. 1, pp. 77 – 84, 2007.

[4] Z. Wang and H. Gu, "A bristle-based pipeline robot for ill-constraint pipes," *IEEE-ASME Transactions on Mechatronics*, vol. 13, no. 3, pp. 383–392, Jun 2008.

[5] J. Qiao, J. Shang, and A. Goldenberg, "Development of Inchworm In-Pipe Robot Based on Self-Locking Mechanism," *IEEE-ASME Transactions on Mechatronics*, vol. 18, no. 2, pp. 799–806, Apr 2013.

[6] M. Meribout, "A Wireless Sensor Network-Based Infrastructure for Real-Time and Online Pipeline Inspection," *IEEE Sensors Journal*, vol. 11, no. 11, pp. 2966–2972, Nov 2011.

[7] A. Cataldo, G. Cannazza, E. De Benedetto, and N. Giaquinto, "A New Method for Detecting Leaks in Underground Water Pipelines," *IEEE Sensors Journal*, vol. 12, no. 6, pp. 1660–1667, Jun 2012.

[8] J.-H. Kim, J.-C. Lee, and Y.-R. Choi, "LAROB: Laser-Guided Underwater Mobile Robot for Reactor Vessel Inspection," *IEEE-ASME Transactions on Mechatronics*, vol. 19, no. 4, pp. 1216–1225, Aug 2014.

[9] J. Nestleroth, "Pipeline in-line inspection challenges to ndt," *insight-wigston then northampton*, vol. 48, no. 9, p. 524, 2006.

[10] R. Amineh, N. Nikolova, J. Reilly, and J. Hare, "Characterization of surface-breaking cracks using one tangential component of magnetic leakage field measurements," *IEEE Transactions on Magnetics*, vol. 44, no. 4, pp. 516–524, April 2008.

[11] P. Laursen, G. C. Vradis, and C. Swiech, "First robotic device to inspect unpiggable gas transmission pipeline," *Pipeline & Gas Journal*, vol. 236, no. 11, 2009.

[12] R. Ireland and C. Torres, "Finite element modelling of a circumferential magnetiser," *Sensors and Actuators A: Physical*, vol. 129, no. 1, pp. 197–202, 2006.

[13] D. Pereira and T. G. R. Clarke, "Modeling and Design Optimization of an Eddy Current Sensor for Superficial and Subsuperficial Crack Detection in Inconel Claddings," *IEEE Sensors Journal*, vol. 15, no. 2, pp. 1287–1292, Feb 2015.

[14] G. Yang, G. Dib, L. Udpa, A. Tamburrino, and S. S. Udpa, "Rotating Field EC-GMR Sensor for Crack Detection at Fastener Site in Layered Structures," *IEEE Sensors Journal*, vol. 15, no. 1, pp. 463–470, Jan 2015.

[15] Y. He, G. Tian, H. Zhang, M. Alamin, A. Simm, and P. Jackson, "Steel Corrosion Characterization Using Pulsed Eddy Current Systems," *IEEE Sensors Journal*, vol. 12, no. 6, Jun 2012.

[16] P.-S. Murvay and I. Silea, "A survey on gas leak detection and localization techniques," *Journal of Loss Prevention in the Process Industries*, vol. 25, no. 6, pp. 966–973, 2012.

[17] C. Hellier, "Handbook of nondestructive evaluation," 2001.

[18] W. Du and S. Yelich, "Post-earthquake pipeline leak detection technologies," *Smart Sensors and Sensing Technology*, pp. 265–283, 2008.

[19] D. Jiles, "Review of magnetic methods for nondestructive evaluation," *NDT international*, vol. 21, no. 5, pp. 311–319, 1988.

[20] H. Baflausen, "Two phenomena revealed with the help of new amplifiers," *Phys Z*, vol. 29, p. 401, 1919.

[21] M. R. Nabavi and S. N. Nihtianov, "Design Strategies for Eddy-Current Displacement Sensor Systems: Review and Recommendations," *IEEE Sensors Journal*, vol. 12, no. 12, pp. 3346–3355, Dec 2012.

[22] J. C. Martins and P. Seleghim, "Assessment of the performance of acoustic and mass balance methods for leak detection in pipelines for transporting liquids," *ASME Journal of Fluids Engineering-transactions*, vol. 132, 2010.

[23] D. Bo, Z. Huiping, S. Sha, and T. Jian, "Research on ultrasonic inspection of pipeline corrosion," in *IEEE International Conference on Control and Automation*, May 2007, pp. 2572–2575.

[24] A. Ferrari, "Modelling approaches to acoustic cavitation in transmission pipelines," *International Journal of Heat and Mass Transfer*, vol. 53, no. 19, pp. 4193–4203, 2010.

[25] C. Frey, "Rotating optical geometry sensor for fresh water pipe inspection," in *IEEE Sensors*, Oct 2008, pp. 337–340.

[26] K. Kim, K. Kim, H. Jung, and H. Chang, "Measurement of defect thickness of the wall thinning defect pipes by lock-in infrared thermography technique," in *International Conference on Experimental Mechanics*, vol. 7522, 2009, pp. 75 224Z–75 224Z–6.

[27] A. Zitouni, L. Beheim, R. Huez, and F. Belloir, "Smart electromagnetic sensor for buried conductive targets identification," *IEEE Sensors Journal*, vol. 6, no. 6, pp. 1580–1591, DEC 2006.

[28] H. Goedecke, "Ultrasonic or MFL inspection: Which technology is better for you?" *Pipeline & Gas Journal*, vol. 230, no. 10, pp. 34–41, October 2003.

[29] A. Khodayari-Rostamabad, J. Reilly, N. Nikolova, J. Hare, and S. Pasha, "Machine learning techniques for the analysis of magnetic flux leakage images in pipeline inspection," *IEEE Transactions on Magnetics*, vol. 45, no. 8, pp. 3073–3084, Aug 2009.

[30] S. Burkett and H. Schempf, "Explorer-II: Wireless self-powered visual and NDE robotic inspection system for live gas distribution mains," Carnegie Mellon University, Tech. Rep. REP-GOV-DOE-051231, 2006.

[31] C. E. Shannon, "Prediction and entropy of printed english," *Bell System Technical Journal*, vol. 30, no. 1, pp. 50–64, 1951.

[32] R. N. McDonough and A. D. Whalen, *Detection of signals in noise*. Academic Press, 1995.

[33] Y. Li, Y. Dong, and G. Sheng, "A signum minimum entropy filter for irregularity detection in optical yarn signals," *Measurement Science and Technology*, vol. 21, no. 3, p. 035104, 2010.

[34] A. Sheinker, N. Salomonski, B. Ginzburg, L. Frumkis, and B.-Z. Kaplan, "Magnetic anomaly detection using entropy filter," *Measurement science and technology*, vol. 19, no. 4, p. 045205, 2008.

[35] A. Rényi, "On measures of entropy and information," Berkeley, CA, pp. 547–561, 1961. [Online]. Available: http://projecteuclid.org/euclid.bsmsp/1200512181

[36] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933.

[37] D. Gu and H. Hu, "Renyi entropy based target tracking in mobile sensor networks," in *World Congress*, vol. 18, no. 1, 2011, pp. 13 558–13 563.

[38] Y. Zhang, P. Candra, G. Wang, and T. Xia, "2-d entropy and short-time fourier transform to leverage gpr data analysis efficiency," 2014.

[39] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[40] J. Neyman and E. S. Pearson, "The testing of statistical hypotheses in relation to probabilities a priori," *Joint Statistical Papers*, pp. 186–202, 1967.

[41] A. R. Webb, *Statistical pattern recognition*. Wiley, 2003.

[42] J. García-Martín, J. Gómez-Gil, and E. Vázquez-Sánchez, "Non-destructive techniques based on eddy current testing," *Sensors*, vol. 11, no. 3, pp. 2525–2565, 2011.

[43] D. H. Hur, M. S. Choi, H.-S. Shim, D. H. Lee, and O. Yoo, "Influence of Signal-to-Noise Ratio on Eddy Current Signals of Cracks in Steam Generator Tubes," *Nuclear Engineering and Technology*, vol. 46, no. 6, pp. 883–888, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1738573315301571

[44] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot modeling and control*. John Wiley & Sons New York, 2006.

[45] R. N. MacDonough and A. D. Whalen, *Detection of Signals in Noise.* Access Online via Elsevier, 1995.

[46] H. Shimazaki and S. Shinomoto, "A method for selecting the bin size of a time histogram," *Neural Computation*, vol. 19, no. 6, pp. 1503–1527, 2007.

PLACE PHOTO HERE

Farid Sheikhi Biography text here.

Davide Spinello Biography text here.

Wail Gueaieb Biography text here.