Article:

1 **Microphone array analysis for simultaneous condition detection, localization and classification in a pipe**

2

3 Yicheng Yu[a,*], Rob Worley[b], Sean Anderson[b], Kirill V. Horoshenkov[a]

4 [a] Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK.

5 [b] Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD,

6 UK

7 * Corresponding author. E-mail address: Yicheng.Yu@sheffield.ac.uk

8

9 **Abstract**

10 An acoustic method for simultaneous condition detection, localization and classification in air-filled

11 pipes is proposed. The contribution of this work is three-fold: (i) a microphone array is used to extend

12 the usable acoustic frequency range to estimate the reflection coefficient from blockages and lateral

13 connections; (ii) a robust regularization method of sparse representation based on wavelets basis

14 function is adapted to reduce the background noise in acoustical data; (iii) the wavelet components are

15 used to localize and classify the condition of the pipe. The microphone array and sparse representation

16 method enhance the acoustical signal reflected from blockages and lateral connections and suppress

17 unwanted higher-order modes. Based on the sparse representation results, higher-level wavelet

18 functions representing the impulse response are used to localize the position of the sensor corresponding

19 to a blockage or lateral connection with higher spatial resolution. It is shown that the wavelet

20 components can be used to train and to test a support vector machine (SVM) classifier for the condition

21 identification more accurately than with a time-domain SVM classifier. This work paves the way for

22 the development of simultaneous condition classification and localization methods to be deployed on

23 autonomous robots working in buried pipes.

24 **I.   Introduction**

25 Buried pipe infrastructure is important to urban life and forms a vital part of many engineering structures

for transporting fluids and gases. In the UK alone there are over 600,000 km of sewer pipes [1]. The US Environmental Protection Agency estimates that water collection systems in the USA have a total replacement value between $1 and $2 trillion. In Europe, buried water pipe networks are much longer and have a much higher replacement value. These networks are aging rapidly and becoming more heavily used due to population growth, increasing demand for water, and climate change, which leads to an increased rate and severity of faults in these pipes. Therefore, reliable techniques for condition monitoring and fault detection are required for the inspection and targeted maintenance of pipe infrastructure.

Autonomous robotic sensing systems working in buried pipes for condition monitoring and fault detection offer the opportunity to capitalise on recent advances in acoustic and ultrasonic sensing techniques [1]. Acoustic methods have been investigated for blockage detection and condition assessment in sewage pipes in the past decades [2]. These methods are a very attractive alternative to traditional visual closed-circuit television (CCTV) inspection methods because they are rapid and highly efficient computationally. Acoustically reflective artefacts including blockages can be localized remotely with respect to the robot position using the time delay of acoustic echoes measured with a microphone [3]. In sewer pipes the power reflection ratio and signal phase measured with the microphone can be used to discriminate between various in-pipe conditions, e.g. blockage, lateral connection or pipe end [4].

Although acoustic methods are well suited for use on an autonomous robotic platform, they are complicated by the multi-modal sound wave propagation in a partially filled sewer pipe [5]. As a result, in this class of applications it is common to limit the frequency range to the so-call plane wave regime only, i.e. to the range below the first eigen-frequency of the round sewer pipe, $f_{10} = \frac{0.59c}{2R}$ [6], where c is the sound speed in air and R is the radius of the pipe. In the case of a typical 300 mm sewer pipe this frequency is 669 Hz for c = 340 m/s. In order to radiate a sufficient acoustic power in this frequency range a large powerful speaker is generally required, which is difficult to deploy on a small robot that would operate in a typical sewer. Furthermore, such a low-frequency range limits the condition localization and classification accuracy due to a relatively long wavelength and restricted frequency

band. The main contribution of this paper is to overcome this limitation by proposing a new microphone array processing and machine learning method that extends the frequency range well above the first eigen-frequency to achieve much higher spatial resolution for condition detection and classification in sewer pipes.

A microphone array consists of a set of microphones positioned in a specific way to capture the spatial information about the sound field that can be used for various purposes, e.g. spatial filtering, noise reduction and dereverberation problems for audio processing [7]. This paper uses the microphone array located at the same cross-section to capture the acoustic signal containing the spatial information about the first four modes. This information is then processed to extract the fundamental mode (plane wave) to enable defect localization and classification. This method makes use of wavelets which are well suited to reconstruct a transient signal in the presence of background noise. The idea of using wavelets to deal with transient signals is motivated by prior work such as Ferrante et al [8] who used the wavelet transform to analyze the transient pressure signal for leakage detection and Owowo and Oyadiji [9] who used wavelets and the soft threshold method to cancel background noise from acoustic signals for leakage detection in an air-filled pipe. However, little or no work has been on the use of wavelets to identify and localize conditions in sewer pipes. This paper proposes a sparse representation method which uses a wavelets basis to cancel the background noise, improve the resolution for condition localization and to increase the accuracy of classification between blockages and lateral connections through post-processing of enhanced acoustical data with Support Vector Machine (SVM).

A health monitoring system should seek to answer a number of key questions, including the presence of a fault (i.e. a blockage or lateral connection) and the location of a fault [10] that are needed to target repair or clean-up operation. An advantage of the microphone sensing array we propose here is that it has a dual use for both condition detection and localization. The method we propose for localization uses the Kalman filter, which operates on acoustic features extracted from the signals measured by the microphone array. This makes the approach highly efficient as a single sensing method is used for both tasks.

The structure of this paper is organised as follows. Section II discusses the theory of acoustic wave

1 propagation in a cylindrical pipe, signal processing theory including wavelets, sparse representation and

2 SVM. Section III presents the simulation results of the microphone array processing and acoustic

3 reflection from blockages and lateral connections in the pipe. The experimental setup is discussed in

4 section IV. Experimental results for blockage localization and identification are discussed in Section V.

5

II.   Theory

II.A. Acoustic waves in a cylindrical pipe
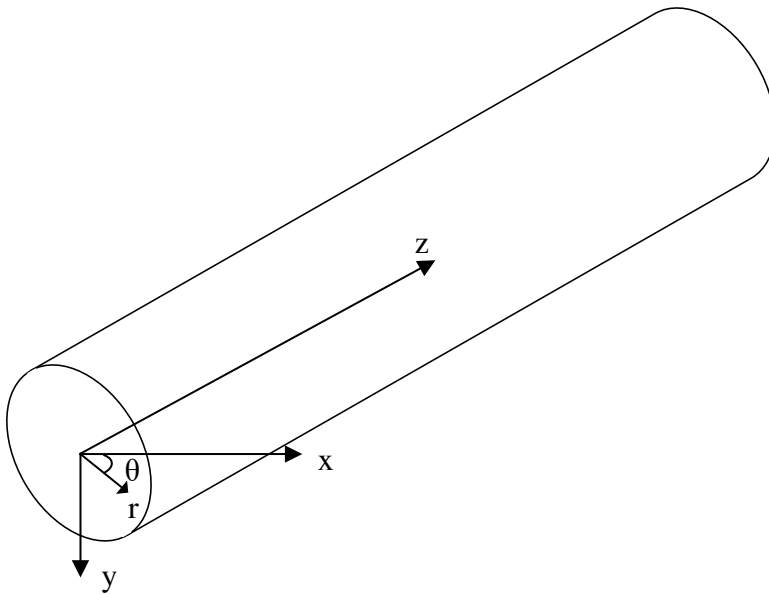
Figure 1. The system of coordinates in a cylindrical pipe

10 The acoustic field in a rigid cylindrical pipe filled with air is the solution of the wave equation written

11 in cylindrical coordinates (r, θ, z) as illustrated in Figure 1. A convenient representation of the total

12 acoustic field in the frequency domain is the normal mode decomposition as suggested by Morse and

13 Ingard [6]:

$$p(r, \theta, z, \omega) = \sum A_{mn} \Psi_{mn}(r, \theta) e^{i \gamma_{mn} z} \tag{1}$$

14 where ω is the angular frequency, *m* and *n* are the mode indices, Ψ_{mn} is the mode shape function of a

15 duct cross-section. The acoustic field in the pipe is multi-modal, i.e. it is a superposition of an infinite

4

1   number of waves with the modal amplitude A_{mn} and coordinate dependent shape functions $\Psi_{mn}(r,\theta)$.

2   For a cylindrical pipe with the radius R the mode shape function is given by [6]:

$$\Psi_{mn}(r,\theta) = \cos(m\theta)J_m(k_{mn}r) \tag{2}$$

3   where $J_m(\cdot)$ denotes the mth Bessel function. The eigen-number γ_{mn} can be obtained from the equation

4   for zero velocity on a rigid wall pipe [6] yielding:

$$J'_{\text{m}}(k_{mn}r)|_{r=R} = 0. \tag{3}$$

5   In the above equation $'$ denotes partial derivative with respect to r. z-axis wavenumber is given by [6]:

$$\gamma_{mn} = \sqrt{k_0^2 - k_{mn}^2}, \tag{4}$$

6   where k_0 is the wavenumber in a free space ($k_0 = \omega/c_0$, c_0 is sound velocity in air).

7   Eq. (4) predicts the wavenumber for different modes at different frequencies, which means that the

8   sound velocity in each mode (except in the case of plane wave when $k_{00} = 0$) is frequency and mode

9   dependent. When the free field wavenumber k_0 is larger than the eigen-number k_{mn}, or the frequency is

10   above the corresponding eigen-frequency f_{mn}, the particular acoustic mode can propagate along the pipe

11   with relatively little attenuation. Figure 2 shows schematically the angular and radial dependence of the

12   first four mode shapes in the cylindrical pipe. In this figure the plus or minus correspond to the sign the

13   modal shape (Ψ_{mn}) in eq. (1) takes.

Mode (0,0)   Mode (1,0)   Mode (2,0)   Mode (0,1)
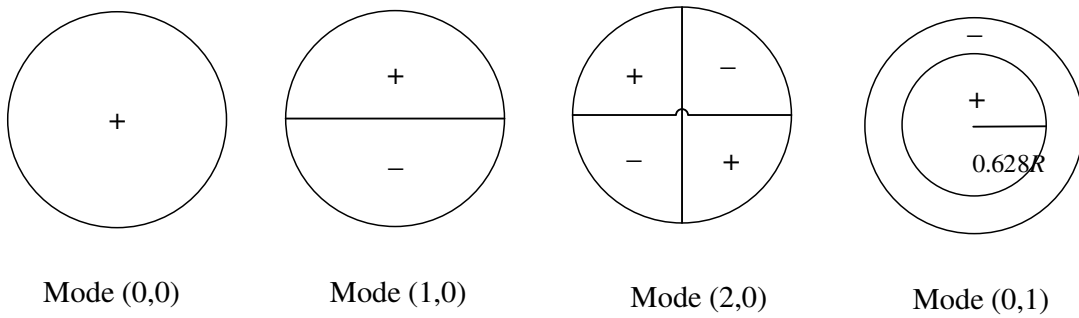
14

15   *Figure 2. An illustration of the behaviour of the first 4 mode shapes in the cylindrical pipe.*

16   The frequency response function (FRF) between the acoustic pressure p at (r,θ,z) and input point

17   source at $(r_s, 0,0)$ with volume velocity Q can be calculated from [6] [11] [5]:

$$\frac{p(\omega)}{Q(\omega)} = \frac{\omega\rho_0}{\pi R_e^2} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{J_m(k_{mn}r_s)\cos m\theta \left[J_m(k_{mn}r)e^{i|z|\gamma_{mn}}\right]}{(\delta_{m0}+1)\gamma_{mn}J_m^2(k_{mn}R)\left[1 - \left(\frac{m}{k_{mn}R}\right)^2\right]}. \tag{5}$$

1

2   In particular, for the point source located at the centre of the pipe cross-section, only the axisymmetric

3   mode can be excited when $m=0$ in Eq. (5). Similarly, only the axisymmetric modes can be separated

4   when the receiver is located at the centre.

5   The sound pressure in an infinite pipe with a blockage or lateral connection is equal to the summation

6   between the incident (p_i) and reflected (p_r) components, given by:

$$p(r,\theta,z,\omega) = p_i + p_r = \sum (A_{mn}e^{ik_z z} + R_{mn}A_{mn}e^{-ik_z z})\Psi_{mn}(r,\theta) \tag{6}$$

7   where R_{mn} is the reflection coefficient from the artefact. Using the modal orthogonality, the acoustic

8   pressure of mode (m,n) can be estimated from [6]:

$$P_{mn}(z) = A_{mn}e^{ik_z z} + R_{mn}A_{mn}e^{-ik_z z} = \iint p(r,\theta,z,\omega)\Psi_{mn}(r,\theta)rdrd\theta/S, \tag{7}$$

9   where S is the cross-sectional area of the pipe. The modal amplitude P_{mn} can be used to predict the

10   absolute value of the acoustic reflection coefficient R_{mn} [6]:

$$R_{mn} = \frac{\max(|P_{mn}(z)|) - \min(|P_{mn}|)}{\max(|P_{mn}(z)|) + \min(|P_{mn}|)} \tag{8}$$

11   Using Eq. (8), the reflection coefficient of a blockage and lateral junction can be predicted via numerical

12   simulations, e.g. Finite Element Method (FEM), to compare against the experimental result as shown

13   in the following sections.

14   Note that in this paper only the plane wave (mode (0,0)) is analyzed for the identification and

15   classification of blockages and lateral connections. This is because the acoustic plane wave is not

16   dispersive, i.e. its speed does not depend on the frequency of sound [6]. As a result, the behavior of the

17   reflection coefficient of the plane wave is easier to interpret and to use for condition classification and

18   localization. For this mode the integral in Eq. (7) can be simplified as:

$$P_{00}(z) = \iint p(r, \theta, z, \omega) r \, dr \, d\theta / S \qquad\qquad (9)$$

1 In Eq. (*9*), the accuracy of the integration depends on the number and distribution of the sensing points

2 and on the choice of the frequency range. In practical situations, there is usually not enough sensing

3 points to cover the whole cross-section over which the integral in Eq. (9) is taken. This is particularly

4 important at high frequencies to lead to unwanted contribution from higher modes in $P_{00}(z)$. In order

5 to remove the unwanted residue of higher modes as well as some background noise from the measured

6 or predicted acoustic pressure, a sparse representation method using wavelets basis function is proposed

7 and studied in this paper.

8 **II.B. Sparse representation using wavelets**

9 Acoustic wave in the pipe can be complicated with multiple modes which travel at different velocities.

10 Using the plane wave mode reconstruction method introduced in Section II.1 (i.e. Eq. (9)) a single mode

11 can be extracted across a broad frequency range and used for in-pipe condition detection, localization

12 and classification. This paper proposes a sparse representation with wavelets for the simplified impulse

13 response to clean up the higher modes residue and to cancel background noise. Different from the

14 wavelet decomposition and soft shrinkage for noise cancelation proposed in Ref. [6], this paper uses

15 sparse wavelet representation to cancel background noise and to clear up some higher modes residue

16 after the plane wave reconstruction with Eq. (9).

17 There are two main reasons for using the sparse representation, Firstly, it is possible to assume that the

18 acoustic echoes from the pipe artefacts (e.g. blockages/junctions) have relatively short duration time.

19 This means that the impulse response actively measured on the robot is considered to contain a large

20 number of zero components apart from the initial pulse and reflected echo wave packs, which leads to

21 the sufficient sparsity relative to its dimension in time domain. The sparse nature of the impulse

22 responses has been illustrated in Ref. [3] as an example and is explained further in our paper. Secondly,

23 the impulse response can be written in terms of appropriate basis vectors where only a few vectors are

24 active, hence reducing the number of the time-domain signals required to store for an accurate signal

25 representation in accordance with the Nyquist sampling theorem [12]. The basis vectors used in this

26 paper are wavelet functions.

Wavelets is an extended Fourier analysis that make use of more general orthogonal bases rather than classical sinusoidal functions. This approach is helpful to overcome the uncertainty principle by exploiting a multi-resolution decomposition [12] to balance between different time and frequency fidelities in different frequency bands. It is particularly useful for decomposing complex signals that arise from multi-scale processes such as images and audio signals [12]. For a given mother wavelet function $\varphi(t)$, the subspace of scale a is generated by the functions (sometimes called child wavelets) [12]:

$$\varphi_{j,q}(t) = a^j \varphi\left(a^j(t - qb)\right) \tag{10}$$

where $j, q \in \mathbb{Z}$; $a > 1$ defines the scale and b is any real number that defines the shift. . j is denoted as the level number of wavelet. In this paper, the typical Daubechies wavelet family was used with $a=2$, and $b=1$
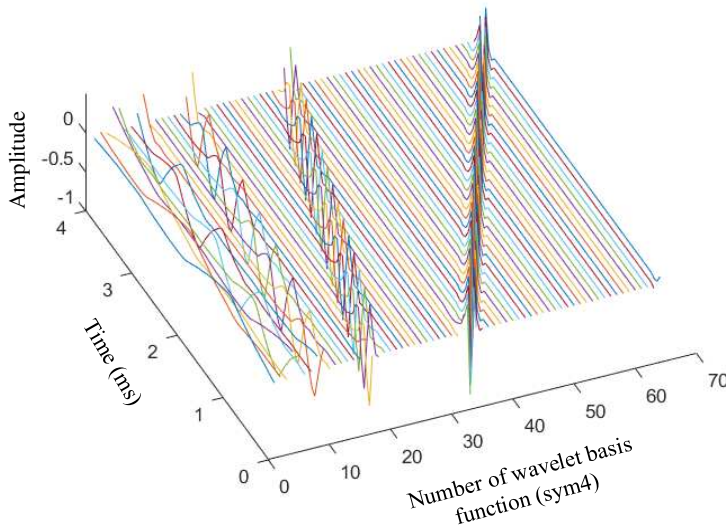
Figure 3. fourth order basis functions of sym4

Acoustic impulse response measured in pipes can be represented accurately using a limited number of wavelet basis functions. An acoustic impulse response of the pipe, $x \in R_n$, may be written as a sparse vector $s \in R_n$ (containing mostly zeros) in the form of a transform basis $\mathbf{W} \in R_{n \times n}$ [12]:

8

$$x = \mathbf{W}s \tag{11}$$

1 As shown in Figure 3, the wavelet matrix $\mathbf{W}$ is generated using *Symlets* wavelet functions, *sym4*, (i.e.

2 generated by the Matlab function *@wmpdictionary*). *Symlets* are modified version of Daubechies

3 wavelets with increased symmetry [12]. The sparsest solution $\hat{s}$ (i.e. the solution represented with a

4 minimum number of wavelet basis functions) satisfies the following optimization problem [12]:

$$\hat{s} = \underset{s}{arg\ min} \|s\|_0 \text{ subject to } \mathbf{W}s - x = 0 \tag{12}$$

5 where $\|.\|_0$ denotes the ℓ_0 pseudo-norm, which is the number of the non-zero components of the vector.

6 This is also referred to as the cardinality of s [12]. Effectively, $\hat{s}$ is the vector which is composed of the

7 amplitudes of the wavelet basis functions at different levels.

8 The optimization problem in Eq. (12) is non-convex and its solution is usually found by using a brute-

9 force search [12] which can be computationally expensive. Fortunately, it is possible to relax the

10 optimization in Eq. (12) to a convex ℓ_1-minimization [12]:

$$\hat{s} = \underset{s}{arg\ min} \|s\|_1 \text{ subject to } \mathbf{W}s - x = 0 \tag{13}$$

11 where $\|.\|_1$ denotes the l_1-norm, which describes the sum of absolute values of the vector.

12 A related convex optimization problem is the following:

$$\hat{s} = \underset{s}{arg\ min} \{\|\mathbf{W}s - x\|_2^2 + \lambda\|s\|_1\} \tag{14}$$

13 where $\lambda > 0$ is a parameter that weights the importance of sparsity. The estimated impulse response

14 can then be represented by:

$$\hat{x} = \mathbf{W}\hat{s} \tag{15}$$

15 In this study the SpaRSA algorithm [11] was used to solve Eq. (14). The SpaRSA algorithm provides a

16 general framework for solving the sparse representation problem [13]. It can be viewed as an accelerated

17 version of the iterative shrinkage thresholding algorithm [14]. The regularization parameter λ is usually

18 set as a specific small constant. The SpaRSA algorithm uses an adaptive continuation technique to

optimize and update the value of λ for a more efficient convergence [11]. Key steps in this algorithm are presented in Table 3 in Appendix I.

Different level of wavelets has different frequency components. Wavelets with higher frequency components can provide a higher spatial resolution to the problem of condition and/or robot localization. For example, the fourth order *Symlets* function, *sym4*, has 5 different levels, where s_1 corresponds to the lower frequency and s_5 to the higher frequency components. Therefore, it is convenient to use the higher level of wavelet domain vector $\hat{s}$ to predict the location of the robot with higher precision. Furthermore, wavelet domain vector $\hat{s}$ can also be used to fit into the SVM trainer to identify the blockage from junctions.

II.C. Robot Localization

After the sparse representation of the impulse response, the higher-level wavelet components can be used to localize the robot position along the pipe. Robot localization is the means by which a robot estimates its position with respect to the surrounding environment. Localization is required for robot control and autonomous navigation, reporting the location of conditions detected in a pipe network, and for mapping unknown parts of the pipe network. Normally, information from sensing of the robot's motion and surroundings is input into a localization function. In typical robotic applications, vision and range-finding sensors such as scanning lidar are popular means of making perceptions, as they are able to acquire a large amount of information from the arbitrary environment. In the pipe environment, however, these sensors are limited in scope and only able to observe nearby artefacts due to the confined space within the cross-section that is very limited compared to a relatively long length of the pipe and scale of the overall pipe network. This limit in scope means that a localization estimate will accumulate uncertainty over time, and that the estimate will likely drift from the true robot position. Acoustic echo sensing is able to perceive more distant artefacts in the pipe environment and offers a means of perception which will not cause an accumulation of uncertainty. In previous study [3], the robot localization has been validated with a speaker and single microphone sensing system using the plane wave below the cut-off frequency. This paper uses the microphone array to extend the frequency range of the signal to localize the robot and artefacts more precisely.

1   Robot localization typically takes a *probabilistic* approach [15] where the uncertainty in each

2   measurement is acknowledged and a localization estimate is the *most likely* value of the robot's *state* in

3   the probability distribution computed over all possible states. Many robot localization approaches are

4   derived from a *Bayes filter*, a mathematical tool that facilitates the incorporation of prior knowledge

5   and measurements to produce a *posterior* estimate. A practical implementation of this is the *Kalman*

6   *filter* described below.

7   The robot is modelled as moving in one dimension along the *z*-axis of a pipe with the position at the

8   time *t* given by:

$$z_t = z_{t-1} + u_t + w_t \qquad (16)$$

9   where u_t is the command robot motion and w_t is additive error in the motion, drawn from a normal

10   distribution with standard deviation $\Sigma_{u,t}$. After moving, the robot makes an acoustic echo measurement

11   of the distance to N nearby acoustically reflective artefacts, ξ_t^n, described by:

$$\xi_t = \{\xi_t^1, \xi_t^2, \dots, \xi_t^n, \dots, \xi_t^N\} \qquad (17)$$

12   These measurements correspond to artefacts in the pipe environment the positions of which are to be

13   estimated along with the robot position. Therefore, the *state* is given by:

$$\mathbf{z}_t = [z_t, z^1, z^2, \dots, z^m, \dots, z^M]^T \qquad (18)$$

14   The state space model is therefore given by:

$$\mathbf{z}_t = \mathbf{A}\mathbf{z}_{t-1} + \mathbf{B}(u_t + w_t) \qquad (19)$$

15   where **A** is the identity matrix and $\mathbf{B} = [1, 0, 0, \dots, 0]^T$.

16   The Kalman filter estimate is parameterised as a multivariable Gaussian distribution described by the

17   mean vector $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma}_t$. The *prediction* step of the filter incorporates the motion

18   information. The predicted mean and covariance are given by:

$$\bar{\boldsymbol{\mu}}_t = \mathbf{A}\boldsymbol{\mu}_{t-1} + \mathbf{B}u_t \qquad (20)$$

$$\bar{\boldsymbol{\Sigma}}_t = \mathbf{A}\boldsymbol{\Sigma}_{t-1}\mathbf{A}^T + \boldsymbol{\Sigma}_{u,t} \qquad (21)$$

11

1   where $\Sigma_{u,t}$ is the motion uncertainty.

2   The acoustic echoes are only used to measure the distance to some artefacts in the pipe environment,

3   and so *data association* needs to be computed between echo measurements ξ_t^n and artefacts in the

4   environment z^m. The negative log-likelihood $\pi_{n,m,t}$ of a match is found for each combination of

5   measurement and artefact, by computing:

$$\boldsymbol{\Psi}_{n,m,t} = \boldsymbol{C}_m\overline{\boldsymbol{\Sigma}}_t\boldsymbol{C}_m{}^T + \boldsymbol{\Sigma}_{\xi,t} \tag{22}$$

$$\pi_{n,m,t} = (\xi_t^n - \bar{\xi}_t^m)^2\boldsymbol{\Psi}_{n,m,t}{}^{-1} \tag{23}$$

6   where C_m is the state space output vector which describes the relative direction of the robot and artefact

7   m, $\Sigma_{\xi,t}$ is the measurement uncertainty and ξ_t^m is the expected measurement for the artefact m. The

8   smallest value of $\pi_{n,m,t}$ is found, which corresponds to the most likely match between measurement and

9   artefact. If the value is larger than a threshold for negative log-likelihood $\prod$, then it is likely that the

10 measurement corresponds to a previously unobserved artefact, so a new artefact is added to the estimate

11 ξ_t. If the value is smaller than the threshold, it is likely that the measurement n corresponds to artefact

12 m.

13 Once the data association has been computed, the *update* step of the Kalman filter provides an estimate

14 and uncertainty using the following equations:

$$\boldsymbol{K}_{n,m,t} = \overline{\boldsymbol{\Sigma}}_t\boldsymbol{C}_m{}^T\boldsymbol{\Psi}_{n,m,t}{}^{-1} \tag{24}$$

$$\boldsymbol{\mu}_t = \overline{\boldsymbol{\mu}}_t + \boldsymbol{K}_{n,m,t}(\xi_t^n - \bar{\xi}_t^m) \tag{25}$$

$$\boldsymbol{\Sigma}_t = (\boldsymbol{I} - \boldsymbol{K}_{n,m,t}\boldsymbol{C}_m)\overline{\boldsymbol{\Sigma}}_t \tag{26}$$

15 These two steps, *prediction* and *update*, are computed at each time point t to recursively estimate the

16 robot position and position of surrounding artefacts using the previous estimate and new information.

17 This estimation process can be improved through improved acoustic sensing and processing described

18 in this paper. When high level wavelet components are used, as described in Section II.B, the

19 measurement uncertainty, $\Sigma_{\xi,t}$, and subsequent estimate uncertainty, Σ_t, will be reduced, and the

20 likelihood of correct *data association* will be higher, improving robustness. If classification of each

1   artefact in the environment is possible, as described in Section II.D, it can be incorporated into this

2   estimation of data association. Again, this can improve the robustness of the estimation process.

3   **II.D. Support Vector Machine (SVM) classifier**

4   SVM is a special type of feed-forward neural network used in machine learning. Given a set of training

5   samples consisting of pairs of a co-occurrence feature vector and class of pipe defect labels as (X_m,

6   Y_m), in which the subscript m indexes the lengths of the vectors ($Y_m \in \{1, -1\}$, $X_m \in R^n$), the training

7   approach of SVM is to adjust the weights α and biases b ($b \in R$) to search for an optimal hyperplane

8   and maximum margin. The latter is defined as the distance of the closest vectors in both classes to the

9   hyperplane. In this paper, two different data sets were used to train and test the binary SVM classifier:

10   (i) the time domain impulse response from the echo (i.e. $\hat{x}$ in Eq.(15)) is used straightforwardly as X_m;

11   (ii) the wavelet domain data $\hat{s}$ (in Eq.(15)) associated with the pipe artefacts (e.g. blockage/junctions)

12   is applied as the training and testing input. To find the hyperplane, the following quadratic programming

13   problem was solved [10]:

$$\min_{\alpha, b} \phi(\alpha) = \frac{1}{2} ||\alpha||^2 \qquad\qquad (27)$$

$$s.t. \quad Y_m[\alpha \cdot X_m + b] \geq 1 \quad for\ all\ m = 1, \dots, N$$

14   However, it is difficult to straightforwardly solve Eq. (27). Therefore, the Lagrange function was

15   introduced:

$$\max_{\alpha} \min_{\alpha, b} L(\alpha, b, \lambda) == \frac{1}{2} ||\alpha||^2 - \sum \lambda_m \{Y_m (X_m \cdot \alpha + b) - 1\}, \qquad (28)$$

16   where λ_m is Lagrange multipliers. A maximum of the Lagrange function L must satisfy the following

17   conditions by applying a first-order derivation to Eq. (28) respective to b and α as:

$$\frac{\partial L(\alpha, b, \lambda)}{\partial b} = \sum \lambda_m Y_m = 0 \qquad \lambda_m > 0 \qquad (29)$$

18

$$\frac{\partial L(\alpha, b, \lambda)}{\partial \alpha} = \alpha - \sum Y_m \lambda_m X_m = 0 \qquad \lambda_m > 0 \qquad (30)$$

13

The patterns $\boldsymbol{X}_m$ for which $\lambda_m > 0$ are also called support vectors (SVs) which lie exactly on the margins and all the remaining training samples satisfy automatically their constraints (Eq. (27)). Substituting the conditions for Eqs. (29)(30) into the Lagrange function (Eq. (28)), the following dual form is obtained:

$$\max L_p(\boldsymbol{\alpha}, b, \boldsymbol{\lambda}) = \sum \lambda_m - \frac{1}{2}\sum Y_m Y_n \lambda_m \lambda_n \langle \boldsymbol{X}_m \cdot \boldsymbol{X}_n \rangle \tag{31}$$

$$s.t. \ \sum \lambda_m Y_n = 0; \ \lambda_n \geq 0 \tag{32}$$

From Eqs. (31)(32), the optimal solution for λ_m can be determined. In pattern recognition, a decision function which correctly classifies the labelled samples $(\boldsymbol{X}_m, Y_m)$ is defined as:

$$f_{\boldsymbol{\alpha},b}(X) = \text{sgn}(\langle \boldsymbol{\alpha} \cdot \boldsymbol{X} \rangle + \boldsymbol{b}) \tag{33}$$

where:

$$\boldsymbol{\alpha} = \sum \lambda_m Y_m \boldsymbol{X}_m; \ \ \boldsymbol{b} = Y_{\boldsymbol{m}} - \sum \lambda_m Y_m \langle \boldsymbol{X}_m \cdot \boldsymbol{X}_m \rangle \tag{34}$$

To make SVM appropriate for real-world decision, radial basis function (RBF) and polynomial kernels are used to nonlinearly map the input data, X_m:

$$\text{RBF:} \ \ \ K(\boldsymbol{X}, \boldsymbol{X}_m) = \exp\left(-||\boldsymbol{X} - \boldsymbol{X}_m||^2 / 2\sigma^2\right) \tag{35}$$

$$\text{Polynomial:} \ \ \ \ \ \ K(\boldsymbol{X}, \boldsymbol{X}_m) = (\boldsymbol{X}^T \cdot \boldsymbol{X}_m + \beta)^d \tag{36}$$

This paper focuses on the binary classification of blockage in sewer pipes, namely, to identify the acoustic echo from a blockage and to discriminate it from other pipe artefacts. For a complete binary classification, a class can be labelled as either $+1$ or -1 for Y_m. On the other hand, the wavelet components of the artefacts wave pack can be used as the training and testing input X_i for the SVM classifier. To illustrate the advantage of a wavelet, the time domain artefacts wave pack, X_m, was obtained through the following process:

1. Acquisition of impulse response: Speaker sent a chirp signal and simultaneously recorded the response using six microphones. After deconvolution and band-pass filtering (200-3000 Hz), the six channels impulse response was obtained x^m where m=1:6.

1   2. Plane wave construction: Averaging the six-channel impulse response provided the pre-

2   processed plane wave impulse response $\boldsymbol{x}$.

3   3. Denoising and feature extraction using wavelets and sparse representation: After generating the

4   wavelet matrix $\boldsymbol{W}$ using *sysm4* level-5 wavelets, plane wave impulse response $\boldsymbol{x}$ was

5   constructed using sparse representation algorithm (SpaRSA) to obtain the denoised signal $\hat{\boldsymbol{x}} =$

6        $\boldsymbol{W}\hat{\boldsymbol{s}}$ and wavelet components $\hat{\boldsymbol{s}} = [\hat{\boldsymbol{s}}_1\ \hat{\boldsymbol{s}}_2\ \hat{\boldsymbol{s}}_3\ \hat{\boldsymbol{s}}_4\ \hat{\boldsymbol{s}}_5]^T$.

7   4. Localisation: The higher-level wavelet components $[\hat{\boldsymbol{s}}_3\ \hat{\boldsymbol{s}}_4\ \hat{\boldsymbol{s}}_5]^T$ were used to represent the

8   higher frequency signal $\hat{\boldsymbol{x}}_h = \boldsymbol{W}[\hat{\boldsymbol{s}}_3\ \hat{\boldsymbol{s}}_4\ \hat{\boldsymbol{s}}_5]^T$ and to apply the Hilbert transform to $\hat{\boldsymbol{x}}_h$, where

9   the coordinates of the peaks of the envelope were associated with the location of pipe artefacts

10   referred to the robot. For sequential robotic localization, the measured coordinates were then

11   be imported to the ξ_t for Kalman filter to predict the locations.

12   5. Classification: The wavelet components ($\hat{\boldsymbol{s}}$) associated with the artefacts as the input (X_m) and

13   the label Y_m (blockage uses 1 and non-blockage uses -1) were used for SVM training and testing.

14   Figure 4 shows the processing steps for denoising, localization and classification.

15
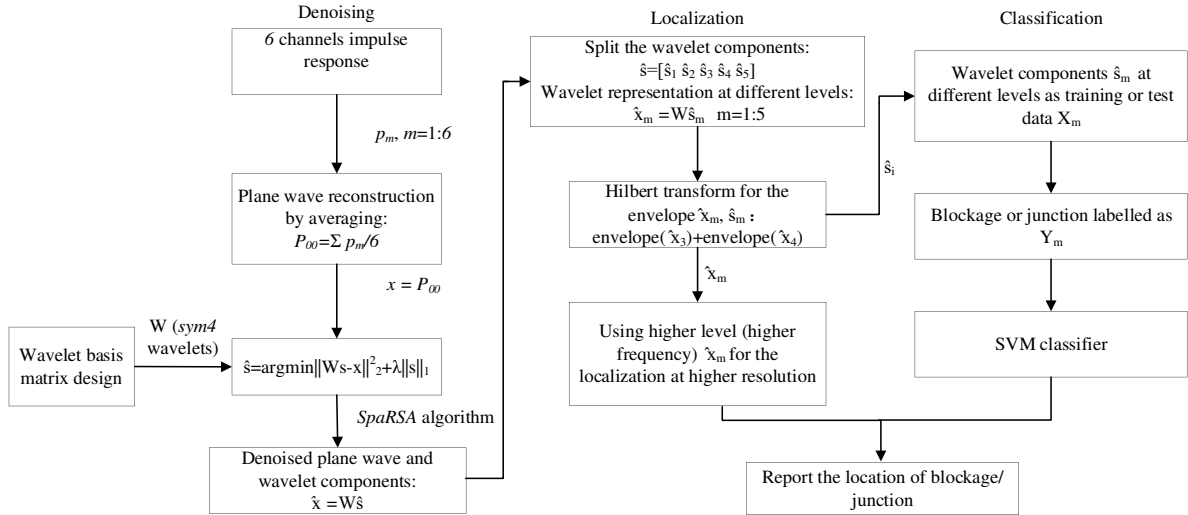
16

17   *Figure 4. Diagram of the process for condition detection and localization*

18 **III.   Simulations**

19   This section discusses the analytical and numerical simulations of the microphone array processing used

to extract the plane wave from the overall acoustic pressure (Eq. (1)) measured on the microphone array and to estimate the reflection coefficient for an artefact in the pipe. The sensor placement and position uncertainties due to the robot movement in the pipe is discussed in section III.1. This can provide evidence in support of the adopted sensor placement strategy for the plane wave reconstruction. The reflection coefficient from blockage and lateral connections will be obtained from numerical simulations to validate the plane wave reconstruction method proposed in the paper via comparison with the experimental results in the following section.

III.A.   Microphone array for plane wave reconstruction

A numerical simulation was implemented based on the transfer function described by Eq. (5). The excitation point source was located close to the pipe wall so that all the acoustic modes were excited. The setup used in this numerical simulation is shown in Figure 5(a). The six virtual microphones were positioned circumferentially and equidistantly spaced at 0.628R. At this radial position $J_0(k_{01}r) = 0$, so that the amplitude of the first axisymmetric mode is equal to zero as illustrated in Figure 2. However, the microphones may not be at the ideal positions in a practical situation, e.g. when the robot platform cannot be perfectly located. Therefore, a simulation using slightly shifted microphones (the centres of the six microphones were shifted at a distance of 0.02R) was also implemented (see Figure 5(b)).

As shown in Figure 5(c), averaging the acoustic sound pressures predicted for the six-microphone array removes the higher modes over the frequency range of 0 - 5kHz (within 1dB fluctuation for plane wave mode) if the microphones were ideally positioned circumferentially at 0.628R. When the sensors shifted slightly at the distance of 0.02R, the first four modes were cancelled significantly over the frequency range 0 - 3.7kHz (within 1.5dB fluctuation) apart from the first eigen-frequency around 1.3kHz. At higher frequencies, the spatial information collected by six microphones tends to be more sensitive to the shifted distance, resulting in a more significant error in the plane wave reconstruction. In this case the fluctuation in the mode (4, 0) and mode (5, 0) is larger than 3dB as shown in Figure 5. Furthermore, it is also observed from additional simulations that the plane wave reconstruction error tends to increase with the microphone shift distance, although the dependence of the error on the microphone shift distance it is not discussed in detail in this paper.
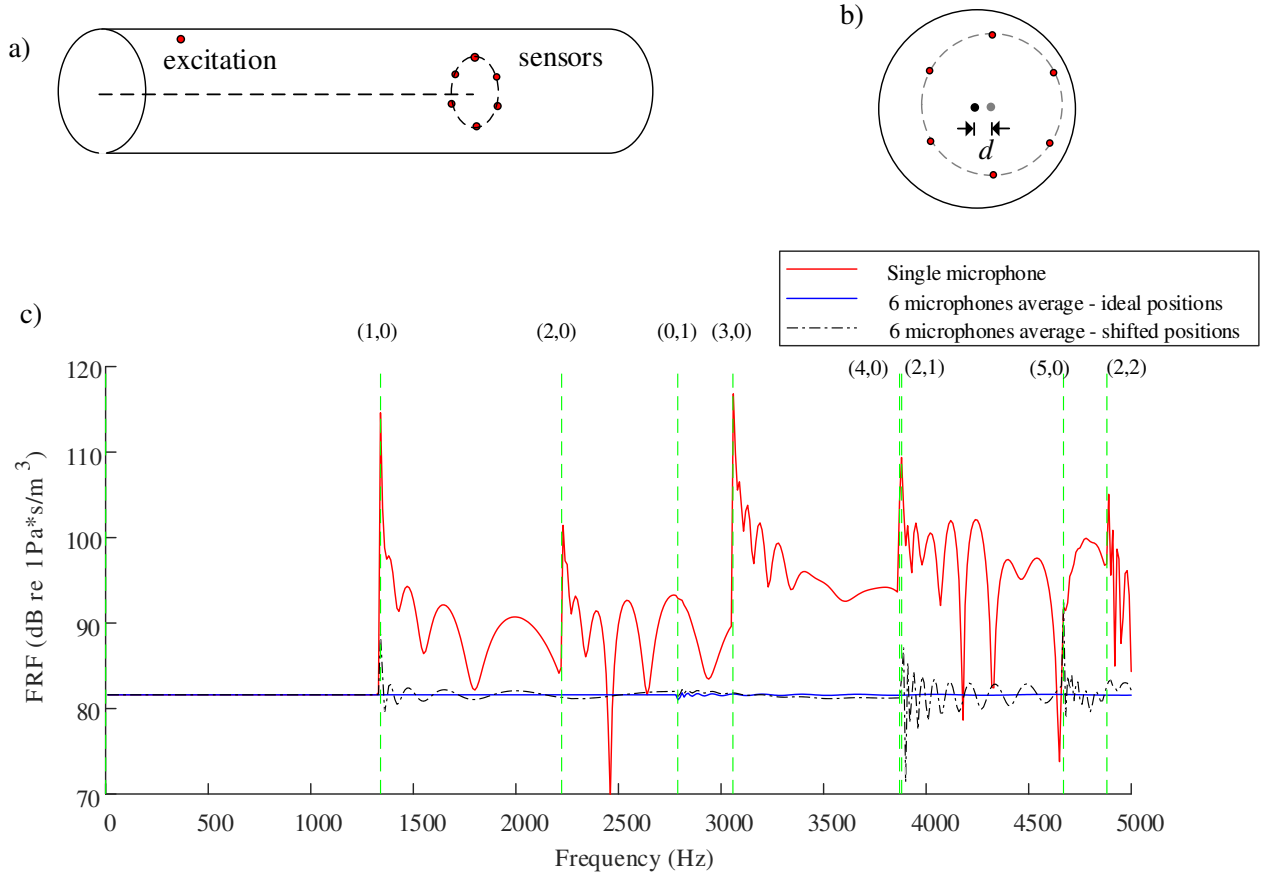
Figure 5. (a) a diagram of the simulation setup with the point source excitation and 6 microphone array; (b) a diagram explaining the simulated microphone array shift; (c) a comparison of the frequency response function between a single microphone and averaged from the six microphones estimated at the ideal or shifted positions

III.B.   Reflection coefficient

In this paper, the acoustic wave reflection from a blockage or lateral junction was studied using the finite element method (FEM) available in commercial software COMSOL. Figure 6 shows the simulation setup for sound propagation in the presence of a blockage and lateral connection. The pipe diameter was 0.15m which is consistent with that used in the experiments. The height of the blockage was set as 0.6 times the pipe diameter, i.e. h/R=1.2 and it was a diameter long (see Figure 5(a)). The maximum mesh size in this numerical study was below 9.5mm which corresponded to 1/12 of the acoustic wavelength at 3kHz. Plane wave excitation was used in the simulations. A Perfectly Matched Layer (PML) was set up at the ends of the pipe to absorb sound to simulate an infinite pipe length. The

17

1   surface of the blockage in this study was assumed solid, i.e. its acoustic characteristic impedance was

2   much larger than that of air ($Z_{blockage} \gg Z_{air}$). The pipe wall was also assumed as rigid.

3

4

5
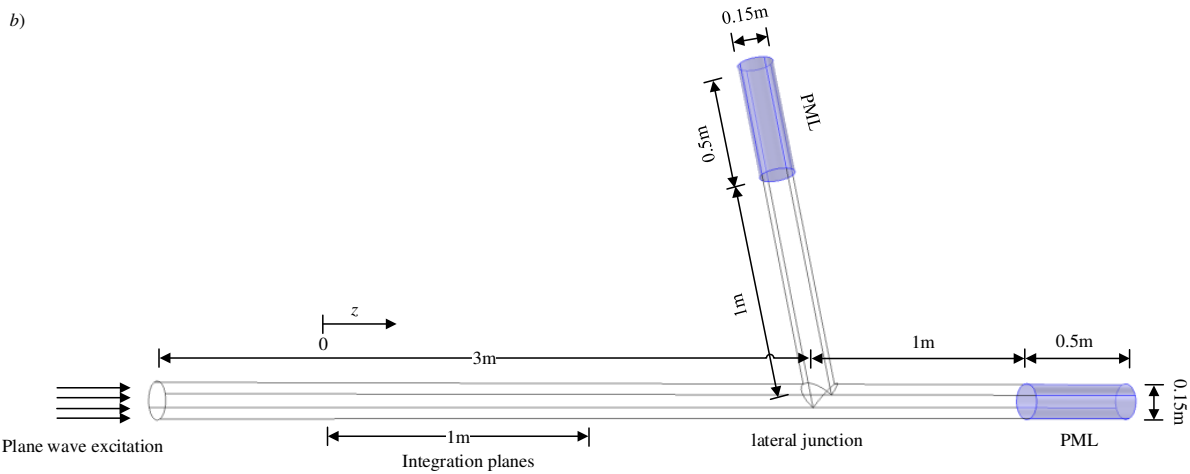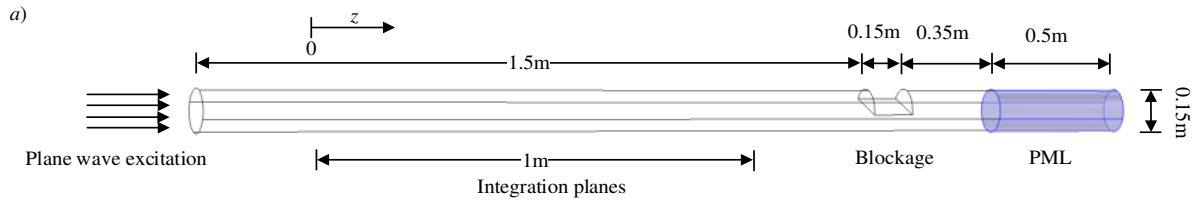
6   *Figure 6. Illustration of the simulation using FE modelling for a pipe with: a) a blockage (h/R=1.2);*

7   *and (b) a 90° lateral connection*

8   As discussed in Section II.1, the sound pressure in the plane wave mode was predicted by Eq. (9) using

9   2D integration over the cross-section. Since the incident and reflected plane wave interferes with each

10 other resulting in the fluctuation of sound amplitude in z-axis. Using the peak and trough value of the

11 fluctuating acoustic pressure at different axial coordinates, the reflection coefficient can be estimated

12 from Eq. (8). The 2D integration over the cross-section was implemented repeatedly with 0.005m

13 intervals and over 1m range.

14 The sound pressure in the plane wave mode, P_{00}, for three frequencies obtained through the simulation

15 and integration in Eq. (9) is shown in Figure 7. These frequencies were chosen to be between the eigen-

16 frequencies and to illustrate the dependence of the sound pressure as a function of the axial direction.

18

1   Note that the curves shown in Figure 7 can be used to determine the amplitude and phase of the complex

2   reflection coefficient for the plane wave mode at these particular frequencies, although only the

3   amplitude of the reflection coefficient is discussed in this paper (see Eqs. (6-9)).

4   It is also worth noting that, although the incident wave was a plane wave, the reflection contains higher

5   modes due to the wave scattering at the artefacts (see Figure 6). There is also a complex relation between

6   the mode number and modal excitation coefficients depending on the nature of an artefact. Using the

7   integral from Eq. (9), the higher-order modes can then be cancelled as discussed in Section III.A.

8   Therefore, the amplitude of P_{00} is a combination of the direct and reflected plane waves. It is frequency

9   dependent because of the complexity of Eq. (5) and integral (9). The modal pressure oscillates as a

10   function of z with the period determined by the wavelength. As shown in Figure 7, the amplitude of the

11   plane wave mode at 1kHz (below the first eigen-frequency for the 0.15 m diameter pipe, $f < f_{10}$) over the

12   axial direction is significantly higher than the amplitude of the 2 kHz wave (between the first and second

13   eigenfrequency: $f_{10} < f < f_{20}$). This can be understood intuitively that a part of the plane wave excitation

14   energy is converted into the higher modes after scattering from the blockage.
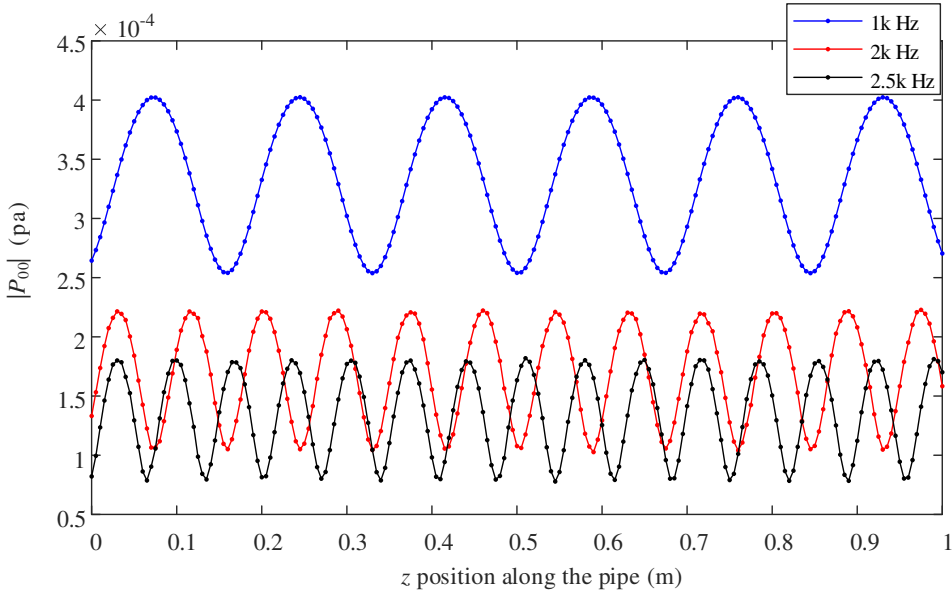
16   *Figure 7. The absolute value of the acoustic pressure of the plane wave mode at three different*

17   *frequencies as a function of the axial distance in the pipe with a blockage (h/R=1.2)*

18   The acoustic reflection coefficient can be calculated from Eq. (8) using the predicted spectra of P_{00}.

19

1   The reflection coefficient simulation results are discussed in the Section V and compared against the

2   experimental data.

IV.   Experimental setup

IV.A.   Robotic platform

5   The acoustic sensing system used in this work consisted of a loudspeaker, six-microphone array and

6   processor (including power amplifier for loudspeaker, ADC, DAC and Raspberry Pi 4 for data

7   acquisition) as shown in Figure 8. This system was installed on a remotely controlled robot (iRobot

8   Looj 330 by iRobot). The sampling rate was 16kHz. A band-pass filter with the frequency response of

9   200 - 3000Hz was used to reduce noise. A 100 – 4000Hz sweep sine with 10s duration was used as the

10   excitation signal. The speaker and microphone array were located at the centre of the pipe within 5mm

11   positional error initially, although this could change due to the robot movement inside the pipe. The

12   radial coordinates of the microphone array were around 60 mm from the pipe centre (see also Figure

13   5(a)). The microphone type used in this test was MSM321A3729H9CP by MEMSensing Microsystems

14   Co., Ltd., and the speaker (Visaton 2242) size was 32mm diameter driven with a 3W power supply.
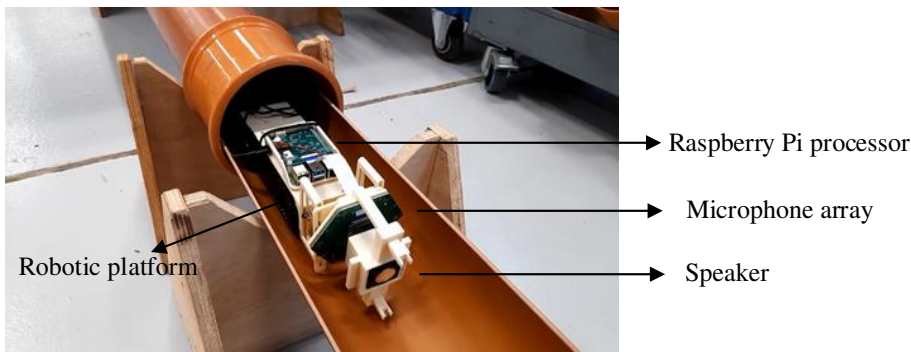
15

16   *Figure 8. Robotic platform and acoustic sensing system*

IV.B.   Pipe network

18   In this work, different sizes of blockages were used in a 150 mm diameter PCV pipe laid in the iCAIR

19   laboratory at the University of Sheffield. These blockages are described by the ratio of the height of the

20   blockage to the pipe radius h/R = 0, 0.4, 0.8, 1.2, 1.6 and 2 as shown in Figure 9(a). Figure 9(b) shows

21   an impression of a sandbag blockage in the 150 mm pipe. Other kinds of blockages were also used in

the experiment, e.g. acoustic absorbent foam and plastic block as shown in Figure 9(c) and (d), respectively. In order to simulate a full 100% blockage, a heavy wooden board was put at the end of the pipe. Efforts were made to seal the circumferential gap between the pipe and board. The straight pipe was constructed from several pipe sections connected with joints at different angles as illustrated in Figure 9(e) and (f). The pipes were not perfectly joined and joints were not perfectly sealed, so some energy in the acoustic wave was able to reflect and leak out due to the discontinuity at a joint.
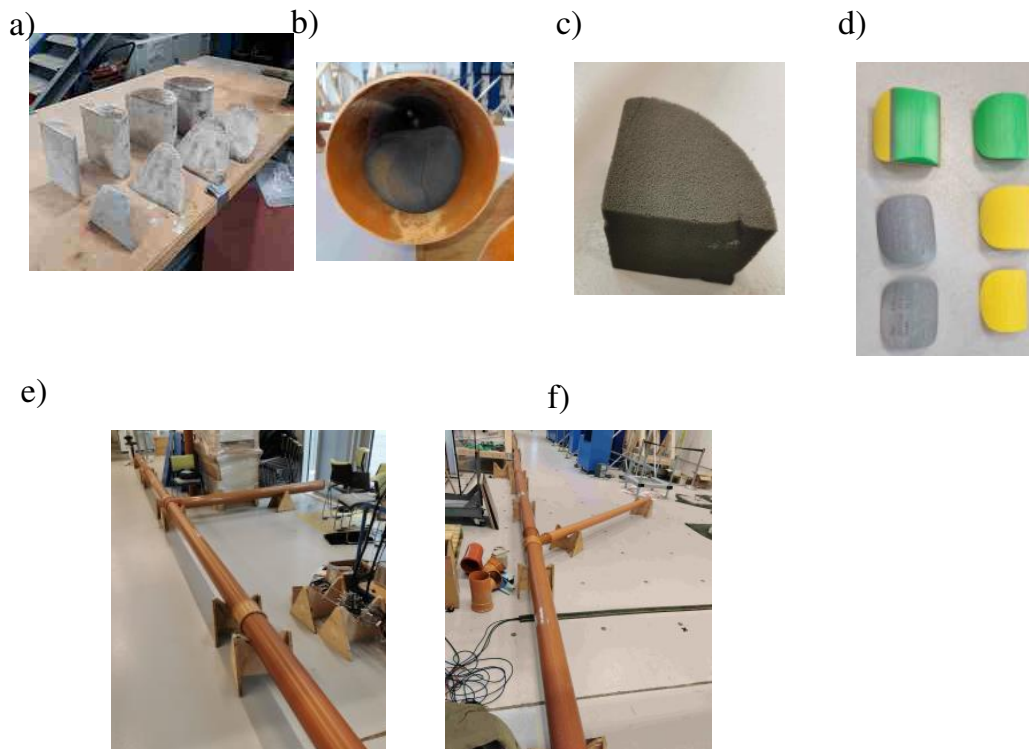
Figure 9. Blockage and lateral junctions simulated in the iCAIR laboratory: (a) different size of concrete blockages; (b) sandbag blockage in the pipe; (c) foam blockage; (d) plastic blockages 3D printed; (e) 90-degree lateral connection; (e) 45/135-degree lateral connection.

V.   Results

V.A. Reconstruction of plane wave and denoising using sparse representation

The impulse response measured using the acoustic system in the pipe with a blockage and lateral connection is shown in Figure 10 and Figure 11, respectively. Note that time domain impulse response was converted into the distance domain response by multiplying the time by sound velocity (343m/s).

1   For a single microphone, the wave dispersion into the higher-order modes significantly complicates the

2   impulse response causing high frequency noise in the data. This noise can cause difficulties in

3   identifying of the condition, particularly when then the condition is a small blockage, e.g. $h/R=0.2$ as

4   shown in Figure 10. Averaging the six-microphone data removes the higher-order modes and provides

5   a cleaner signal than that obtained on a single microphone. This is consistent with the theoretical study

6   presented in Section II.1. Sparse representation cleans up the data further making it more convenient to

7   apply the localisation and classification algorithms detailed in section II. This is even more evident in

8   the case of the data obtained for the pipe with a lateral connection as illustrated in Figure 11.

9
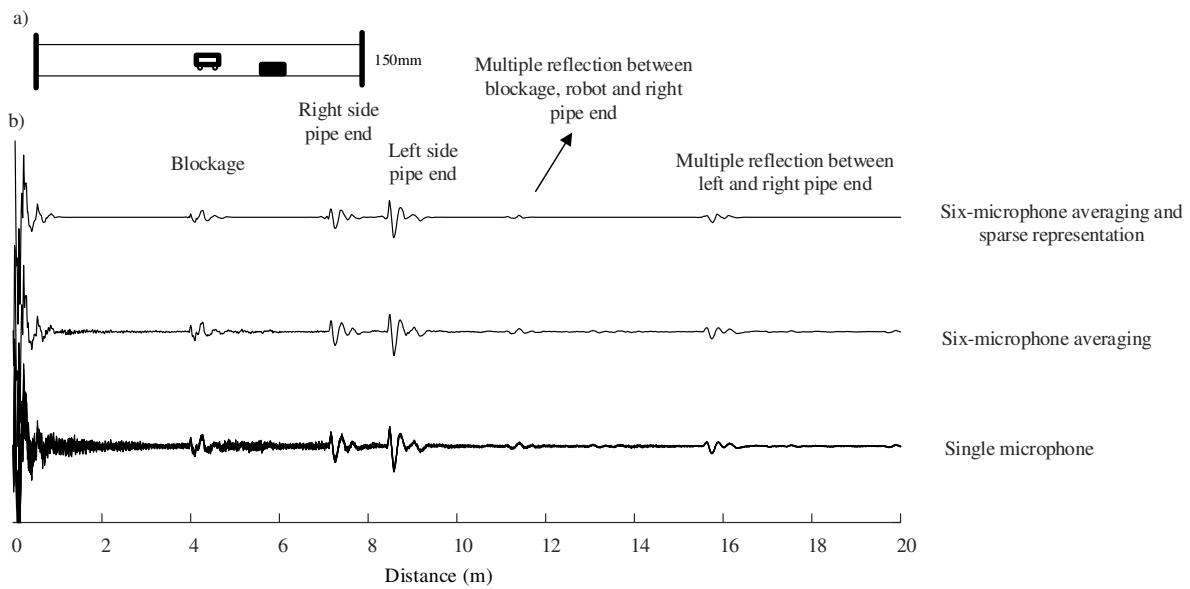
10

11   *Figure 10. Impulse response measured in a straight pipe with blockage (h/R=0.2) located at 4.1m*

12   *away from the robot*
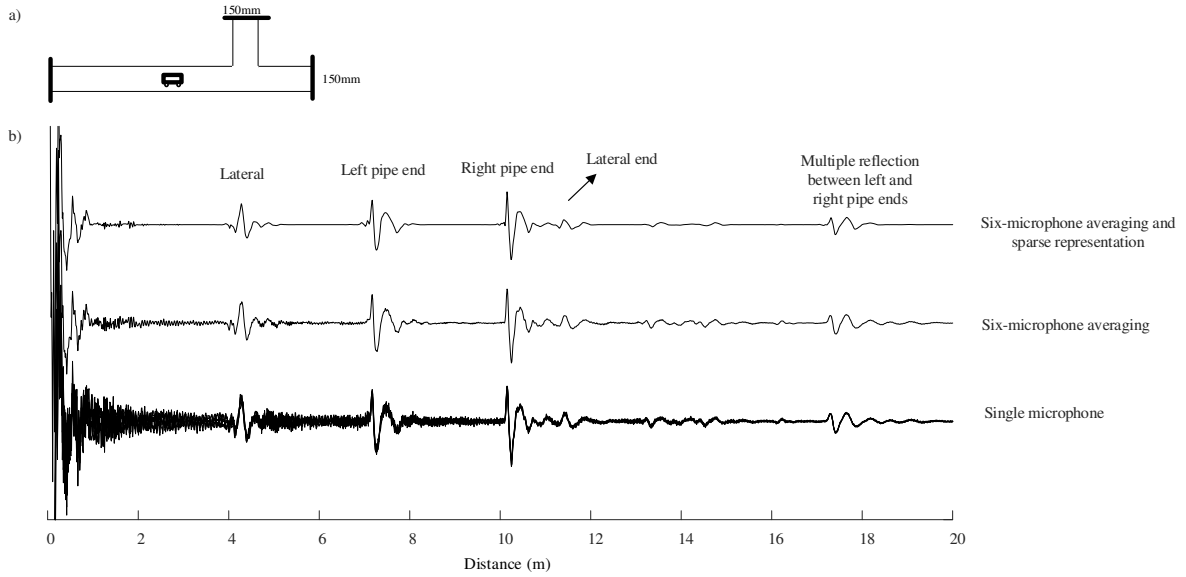
1

2   *Figure 11. Impulse response measured in a straight pipe with 90° lateral connection located at 4.1m*

3   *away from the robot*

4

5   **V.B. Reflection coefficient from blockages and lateral connections**

6   As discussed in Section II.1, the amplitude of the reflection coefficient from a blockage and lateral

7   connection for the plane wave mode can be estimated from Eq. (8). The simulation results for the

8   reflection coefficient were obtained based on the FE modelling in COMSOL (see Section III). For

9   experimental measurements, the impulse response from the blockage and lateral connection (e.g. see

10   the first echo pulse in Figure 10 and Figure 11, respectively) were extracted using time windowing,

11   zero-padded and transferred into the frequency domain. A comparison between the simulation results

12   and measurements of the reflection coefficient spectra from blockage and lateral connection with

13   different setups are shown in Figure 12-14. The measured data were obtained from the denoised plane

14   wave echo using the sparse representation algorithm described in Section V.1.

15   Figure 12 presents the reflection coefficient from blockages with different sizes ($0.2 \leq \frac{h}{2R} \leq 0.8$, see

16   Figure 9(a)) obtained from the simulation and experiments. The full blockage was used as a reference

17   where all the other reflection coefficient were normalized by the full blockage echo obtained

18   experimentally. The simulated results had less than 0.06 average discrepancy with the measurement

23

1   over the frequency range 200-3000Hz. As shown in Figure 12, the acoustic reflection becomes stronger

2   when the blockage size increases. This information can be used to estimate the size of the blockage.

3   Furthermore, the acoustic reflection coefficient becomes larger when the frequency approaches the first

4   cut-off frequency. These properties of the frequency domain spectra can be used to classify blockages

5   from other artefacts such as junctions.

6
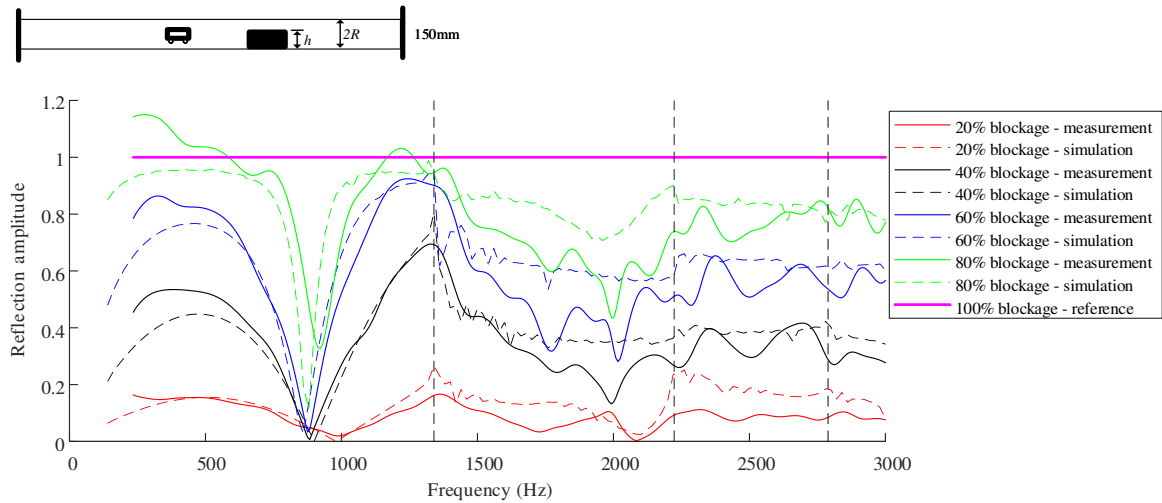
7   *Figure 12. A comparison of the amplitude of the predicted and measured reflection coefficient from a*

8   *blockages of different size (h/2R = 20%, 40%, 60%, 80%, 100%), dashed vertical lines represent the*

9   *cut-off frequencies.*

10   Figure 13 shows the predicted and measured acoustic reflection coefficient spectra from the lateral

11   connection of different diameter attached to the main pipe perpendicularly. The 150 mm diameter

12   branch lateral connection results in a higher reflection coefficient than the 100 mm diameter branch.

13   The lateral connection works like a high-pass filter that allows for the propagation of higher frequency

14   sound waves through the lateral connection where less reflection observed. The reflection coefficient

15   from a lateral connection drops significantly as the frequency of sound approaches the first cut-off

16   frequency. This highlights the importance of extending the frequency range in the proposed analysis to

17   enable to localize and classify conditions beyond a lateral connection.
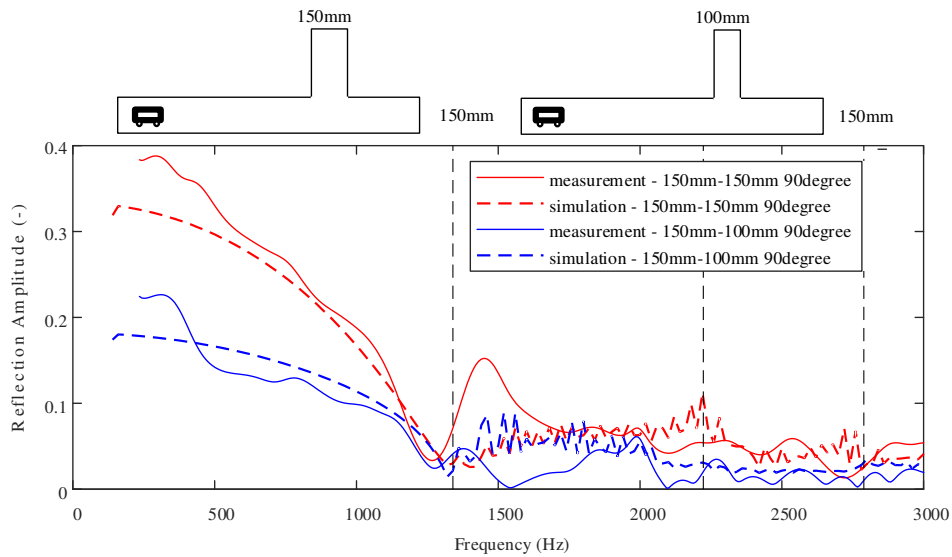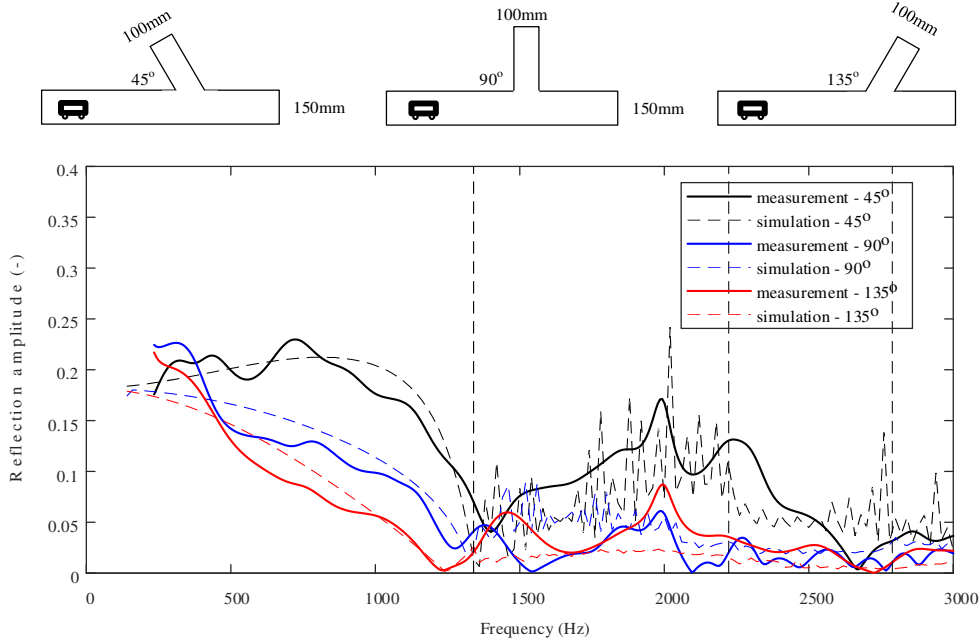
Figure 13. A comparison of the amplitude of reflection coefficient from 90⁰ lateral with different

branch size between the simulation and measurement

Figure 14 shows that the predicted and measured acoustic reflection coefficient from the lateral

connection installed at different angles. An increase in the angle of the lateral connection results reduces

the sound pressure in the reflected plane wave mode over the frequency range 200-3000Hz. For a lateral

connection the reflection coefficient reduces significantly as the frequency of sound approaches the first

cut-off frequency. This slope in the frequency-dependent reflection coefficient becomes steeper when

the lateral angle gets smaller. At higher frequencies beyond the first cut-off frequency, the reflection

coefficient for the lateral connection with an increased angle tends to develop a local peak and beyond

which it decrease gradually. For example, in the case of a 100mm lateral connected at 90⁰ the measured

reflection coefficient increases until reaching a peak around 2kHz then decreases gradually to almost

zero at 2.7kHz (see Figure 13).

Figure 14. Comparison of the amplitude of reflection coefficient from a 45º/135º lateral (100mm branch) with different angle between the simulation and measurement

This section provides the knowledgebase for further identification of blockages and lateral connections according to their acoustic reflection properties. The close agreement (less than 0.06 error on average) between the simulated and measured reflection coefficient for blockages and lateral connections in the frequency domain (300-3000Hz) also validates the signal processing methods of sparse representation. The denoised blockage/lateral signal is useful for the localization and classification algorithms.

**V.C. Robotic localization for blockage/lateral connections**

After the reconstruction of plane wave mode using multiple microphone processing and sparse representation method, a robot can process the denoised signal to localize its position with respect to a blockage or lateral connection. In the previous work [3] the Hilbert transform was used to obtain the envelope of the time domain impulse response where the coordinates of the peaks of the envelope correspond to the relative distance between the pipe artefacts and robot. In this paper, high frequency components of the plane wave impulse response were used for a more precise acoustic localization achieved with microphone array processing. This was accomplished by using the envelope of the higher-level wavelet representation of the impulse response.

As shown in Figure 15(a), the original plane wave impulse response x after the averaging from six-microphone data can be sparse represented and denoised using wavelets to obtain a clearer signal $\hat{x}$ for post-processing. Wavelet components ($\hat{s} = [\hat{s}_1 \ \hat{s}_2 \ \hat{s}_3 \ \hat{s}_4 \ \hat{s}_5]^T$) and their representation ($\hat{x} = W\hat{s}$) for the impulse response are shown in Figure 15(b). Highest-level wavelet components $\hat{s}_5$ are zero in this signal after the shrinkage from sparse representation algorithm. The third and fourth levels of wavelet representation shows higher resolution in the time domain data than the lower levels (see Figure 15). This is because the wavelet components $\hat{s}_3, \hat{s}_4$ correspond to the amplitude of the higher frequency signal. After the representation, $W\hat{s}_3$ and $W\hat{s}_4$ contains the acoustic features wave packs with shorter duration in the time domain, whereas the lower frequency representations, i.e. $W\hat{s}_1$ and $W\hat{s}_2$ provide wider wave packs. Therefore, in this work the higher frequency representations ($W\hat{s}_3 + W\hat{s}_4$) were used to estimate the location of artefacts in the pipe. Specifically, the locations of artefacts were determined using the coordinates of the peaks of the envelope of the signal at high frequencies $W\hat{s}_3 + W\hat{s}_4$. The envelope of the signal $W\hat{s}_3 + W\hat{s}_4$ was calculated using the magnitude of its analytic signal, which was computed by filtering $W\hat{s}_3 + W\hat{s}_4$ with a Hilbert FIR filter of five points length [16] (implemented using the function @*envelope* in Matlab). The envelope results are shown in Figure 16.

The blockage was located at 4m away from the robot with 0.5% prediction error using wavelet representation, whereas the localization error was 4.2% with when the impulse response was used. More experiments were carried out using different size of blockages and lateral connections. The prediction error of higher-level wavelets representation was below 0.7%. This demonstrates an advantage of using wavelets for the robotic localization with higher accuracy and precision.

Although the pipe artefacts can be localized with respect to the position of the robot, the directions of echo pulses are unknown. Sound intensity measurement can be a solution to determine whether the echo comes from front or back of the robot [17]. This paper takes account of the robot position uncertainty into consideration to use the sequential measurement for the localization of artefacts and pipe mapping. The localization measurement using wavelet can be applied to this sequential robotic localization using the Kalman filter as discussed in Section II.3, where the robot moves in the pipe and takes the acoustic measurement sequentially every several meters (i.e. every 2m in this paper).

1
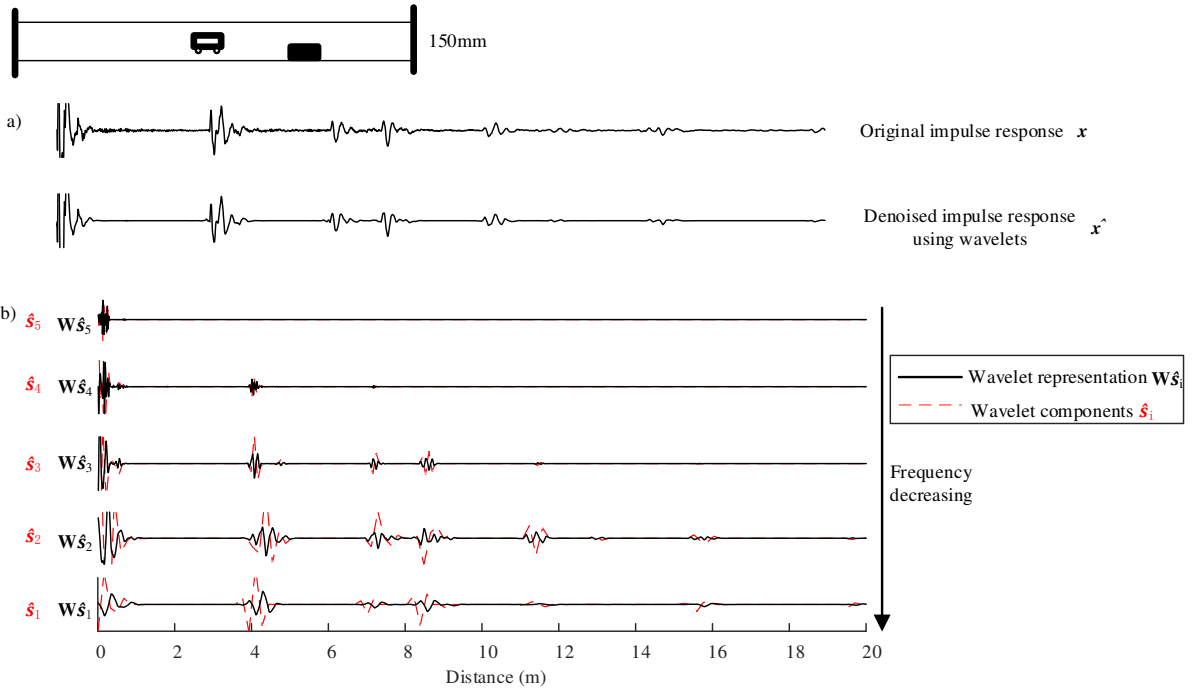
Figure 15. The wavelet components for the impulse response used for the localization in a pipe with a

blockage (h/2R=0.6) at 4m.

4
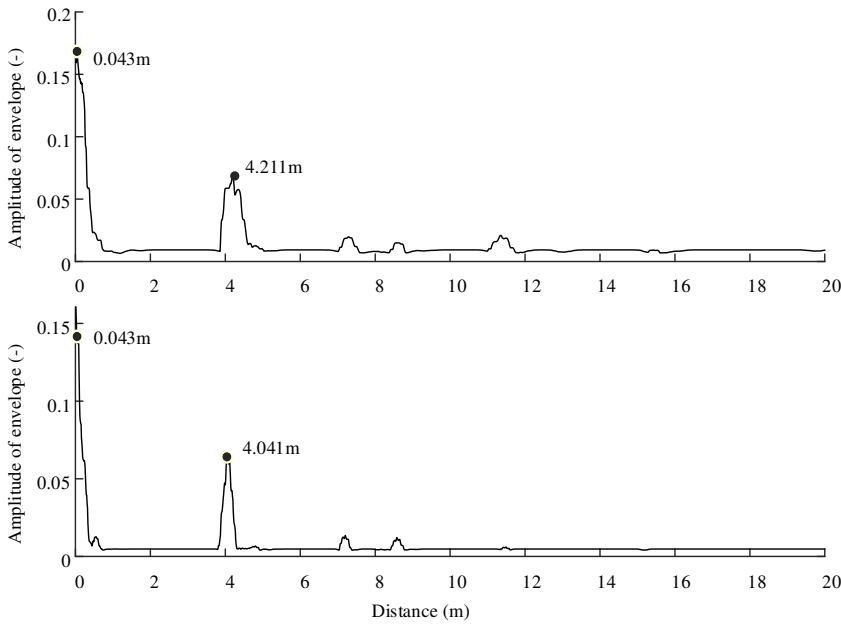
Figure 16. The envelope of: (a) the impulse response; and (b) the wavelet representation at higher-

*level (**W**s₃+**W**s₄)*

| 1 | **V.D. Sequential Robot Localization** |

| 2 | Photographs of the robotic localization test rig are shown in Figure 9 and Figure 17(a) illustrates this |

| 3 | rig schematically. In this experiment a heavy wooden board was installed at the far end of the pipe to |

| 4 | represent a full blockage. The robot was moved towards the lateral connection and stopped every 2m |

| 5 | to measure the impulse response. The measured impulse responses are shown in Figure 17 (b) and (c) |

| 6 | with and without the sparse representation method, respectively. The sparse representation method |

| 7 | removes a significant amount of background noise including dispersive higher modes and some |

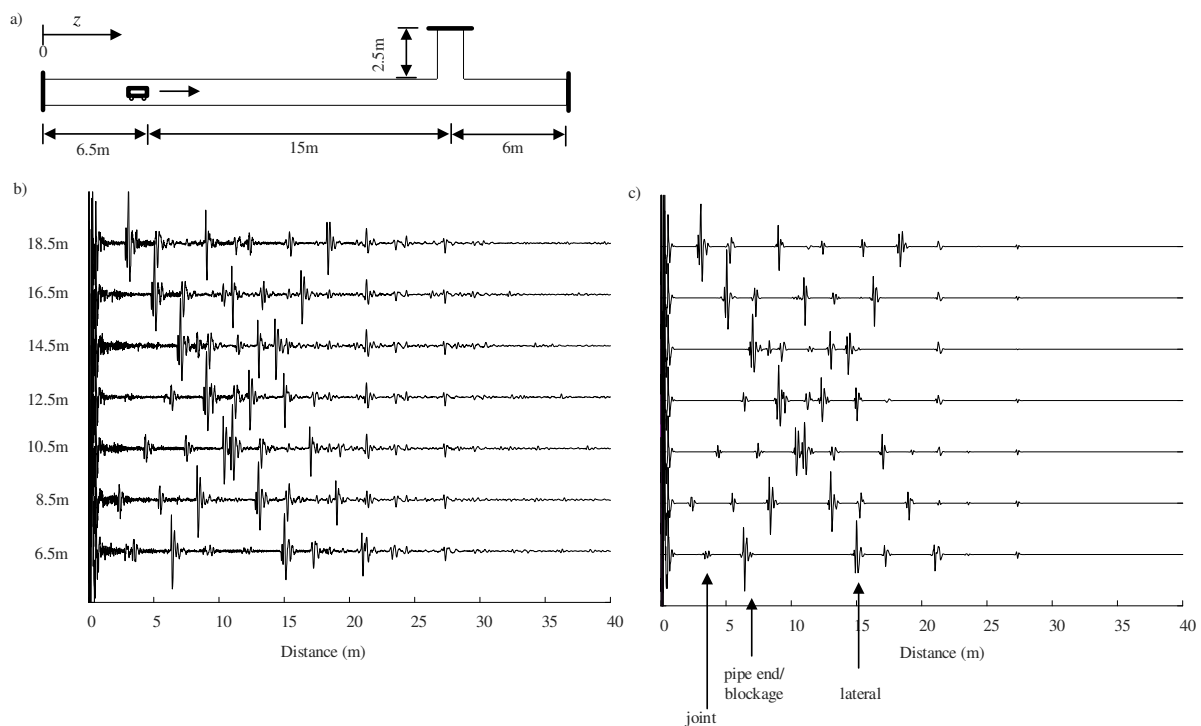| 8 | unwanted reflections from the pipe joints. |

| 9 | |

| 10 | *Figure 17. The robotic sensing system in a pipe network for the localization of blockage/lateral* |

| 11 | *connection: (a) an illustration of the experimental system; (b) 6-microphone average impulse* |

| 12 | *responses when robot was at different positions, without sparse representation processing; (c) the* |

| 13 | *impulse responses after the averaging of 6 microphone data and with sparse representation process* |

| 14 | *when robot was at different positions in the pipe.* |

| 15 | Using the acoustic echoes reflecting from different features in the environment, the robot can estimate |

| 16 | its position while it moves along the pipe. Using the process described in Section II.4, an estimate of |

1 the robot's position can be made by combining the prior estimate of the robot's position with traditional

2 odometry and with new acoustic information obtained from the reported measurements. The uncertainty

3 in new information is incorporated into the estimate, so that the process is designed to have robustness

4 to noise in measurements. However, the measurement noise does have an effect on the precision of the

5 estimate, which is investigated here.

6 Figure 18 shows the results from the simulation of the variation in the mean estimate error for the

7 robot's trajectory (over 100 trajectories) along a pipe obtained for a range of measurement noise levels.

8 This measurement noise was the standard deviation of the Gaussian noise added to each continuous

9 value of distance measurement found from an acoustic echo. For comparison, the estimate made with

10 traditional odometry without using acoustics is also shown. This result illustrates the impact of the

11 uncertainty in robot motion along the pipe. As the acoustic measurement precision increases, the

12 measurement uncertainty decreases, and the median estimate error is seen to decrease from close to the

13 benchmark estimate error of 0.6 metres at 1 metre of measurement uncertainty to 0.25 metres at 0.1

14 metres of measurement uncertainty. This illustrates the strong impact of improved acoustic echo

15 measurement precision on robot localization that can be achieved using the approaches described in this

16 paper.
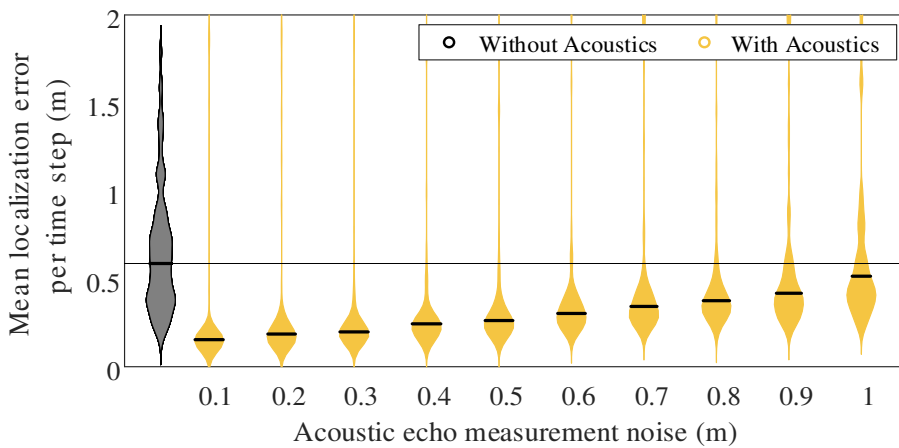
17

18 Figure 18. Localization error per time step with varying measurement noise

19

1   **V.E. Classification of blockages and junctions**

2 In this work, the two classes of pipe artefacts were identified: (i) a blockage; and (ii) a lateral connection.

3 The blockages (35 different types in total) shown in Figure 9 were used for the training and testing of

4 the SVM model. Junctions (25 different types in total) included pipe joints, lateral connections, T-

5 junction and corner junction. As discussed in Section II.4, the time domain wave packs $\hat{x}$ were used

6 directly as input (X_i) for the training and testing. Wavelet components ($\hat{s}$) associated with the artefacts

7 were also used as the input (X_i) where the classifier was expressed as a wavelet SVM classifier. In this

8 study, cross-validation was used with 5 folds, i.e. groups that data samples are split into for the

9 evaluation of SVM modelling to protect against overfitting via data partitioning.

10 For the wavelet based SVM classifier, the accuracy of the blockage detection in front of the robot at the

11 first echo was 88% (53/60) based on the provided cases as shown with the confusion matrix in Table 1.

12 The time domain SVM classifier enabled us to achieve 78% (47/60) accuracy based on the provided

13 cases. It's worth noting that a linear SVM was also implemented in this study with 65% (39/60) and

14 53% (32/60) accuracy using the wavelet components and time domain data, respectively. The detailed

15 accuracy, precision, recall and F1 score [12] for these four classifiers are shown in Table 2. These results

16 provide the evidence that using wavelet components and non-linear kernel (RBF) improves the

17 classification accuracy than using linear kernel and raw time domain data.

18 *Table 1. Confusion matrix of the wavelet SVM classifier from the testing data, TP: true positive, FP: false positive, FN: false*

19 *negative, TN: true negative.*

	Predicted blockage	Predicted junction	
Actual blockage	TP=31	FP=4	35
Actual junction	FN=3	TN=22	25
	34	26	

20 *Table 2. Comparison of time domain linear SVM, wavelet linear SVM, time-domain RBF SVM, and wavelet RBF SVM:*

Metric	Time domain Linear SVM	Wavelet Linear SVM	Time domain RBF SVM	Wavelet RBF SVM

31

Accuracy	53%	65%	78%	88%
Precision	0.571	0.686	0.829	0.886
Recall	0.625	0.615	0.806	0.912
F1 score	0.597	0.649	0.817	0.899

1

2   Eight different settings of pipe network have been used as the demonstration examples, which are shown

3   in Figure 19. In Figure 19, time domain SVM classifier can estimate the first artefacts close to the robot

4   accurately, apart from a small blockage (blockage 1) which presents smaller reflection energy. Whereas,

5   SVM model using the wavelet components as the training and testing data, shows more accurate (10%

6   accuracy improvement, particularly for small blockages) classification result than the time domain

7   SVM classifier. Furthermore, the reflection from joints/lateral connections 7 m behind the robot can

8   also be predicted by wavelet-SVM classifier with around 92% accuracy. However, the time domain

9   SVM enabled us to achieve only 50% prediction accuracy. This provides the evidence that the wavelet

10   method, which takes advantage of the sparsity of the impulse response, can be used to improve the

11   prediction accuracy in comparison to the time domain SVM classification method.
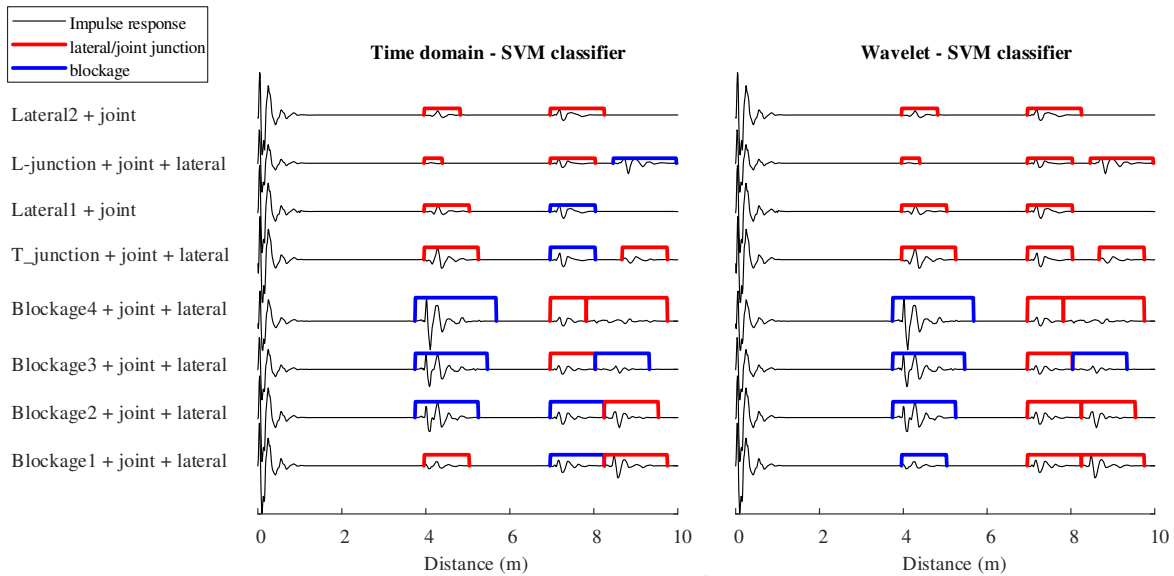
13   *Figure 19. A comparison of the SVM classifier test between time domain impulse response and*

14   *wavelet components*

Although the prediction using wavelet-SVM classifier tends to be accurate in the testing examples (Figure 19), there are some cases in the experiment resulting in error classifications:

1. Small blockages or sound absorbent blockage materials (e.g. acoustic absorption foam in this study see Figure 9(c)). This is because these kinds of blockages do not reflect enough acoustic energy that leads to a negligibly small amplitude of their impulse response and distortion in the assumed reflection coefficient spectra. The smallest successful blockage identified in this work using wavelet-SVM classifier was 20% blockage (blockage 1 in Figure 19).

2. Blockages located close to a junction (<1 m). The acoustic echo from blockage overlaps with the reflection from the junction or another artefact making it difficult to separate multiple reflections and to classify each of them.

3. A robot located too close to the blockage or junction (<1m). This is similar as in case 2 where the multiple reflections occur and overlap.

4. Blockages located behind any artefacts. For example, if the blockage is behind a lateral connection, then the reflected signal can be colored by the presence of this lateral connection. As shown in Figure 19, the lateral connection after the blockage 3 is mistakenly classified.

This classification methods provided in this paper use limited number of blockage and lateral connection cases simulated in the laboratory. More experimental data and realistic environmental testing will be needed to extend this method for multiple classification of different types of blockages and junctions.

VI.   Conclusions

This paper proposed a new acoustic method to simultaneously detect, localize and identify the conditions in an air-filled pipe. Compared with previous studies, the main novel contributions of this paper are: (i) the use of a microphone array to extend the usable acoustic frequency range to estimate the reflection coefficient from blockages and lateral connections; (ii) a robust regularization method of sparse representation based on wavelets basis function adapted to reduce the background noise in the acoustical data; (iii) the use of wavelet components to localize and classify the blockages.

In particular, multiple microphones have been used to reconstruct the plane wave mode beyond the first three eigen-frequencies to support more accurate condition detection, localisation and classification. Numerical and experimental results for the modal reflection coefficient from a blockage and lateral connection have been predicted and compared with measurements. This information has been used to support condition detection and classification.

Wavelet basis functions have been used to sparsely represent the plane wave mode impulse response for the condition detection and classification using the l_1-norm regularization method. The higher-level wavelet functions referring to the higher frequency components of the impulse response have been used to localize the robot and blockage/lateral connection with a higher resolution and accuracy. It has been shown that the wavelet components can also be used to train and to test the SVM classifier for the blockage identification with higher accuracy than using the time-domain SVM classifier.

VII. Acknowledgement

VIII. Appendix I.

Table 3. The Algorithm of Sparse Reconstruction by Separable Approximation (SpaRSA [18]) for l_1-norm regularization

Task: To solve the problem $\hat{\boldsymbol{s}} = \arg\min \frac{1}{2}\|\boldsymbol{W}\boldsymbol{s} - \boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{s}\|_1$
Input: Response signal $\boldsymbol{e}$, wavelet dictionary $\boldsymbol{W}$, parameter $\lambda = 0.001$

Initialization: $k=1$, $\mathbf{A} = \mathbf{W}$, $\pmb{x}_1 = \pmb{x}$, $\tau_1 \mathbf{I} = \mathbf{A}^T \mathbf{A}$, tolerance $\varepsilon = 10^{-5}$
Iteration: 1. $\lambda_k = \max\{0.2\|\mathbf{A}^T \pmb{x}_k\|_\infty, \lambda\}$.
2. Exploit soft shrinkage: $\pmb{s}_{k+1} = \mathrm{shrink}(\pmb{s}_k - \mathbf{A}^T(\mathbf{A}\pmb{s}_k - \pmb{x})/\tau_k, \lambda_k/\tau_k)$ (where $\mathrm{shrink}(s_i, \lambda) = \mathrm{sign}(s_i)\max\{|s_i| - \lambda, 0\}$)
3. Update the step size: $\tau_k = \dfrac{(s_{k+1}-s_k)^T(\nabla\vartheta(s_{k+1})-\nabla\vartheta(s_k))}{(s_{k+1}-s_k)^T(s_{k+1}-s_k)}$
4. If $\dfrac{\|s_{k+1}-s_k\|}{s_k} \leq \varepsilon$, go to step 5. Otherwise, return to step 2
5. $\pmb{s}_{k+1} = \pmb{x} - \mathbf{A}\pmb{s}_{k+1}$
6. If $\lambda_k = \lambda$, stop; Otherwise $k=k+1$, and return to step 1.
Output: $\hat{\pmb{s}} = \pmb{s}_k$, $\hat{\pmb{x}} = \mathbf{W}\hat{\pmb{s}}$

References

[1] "Discover water: treating sewage," 2020. [Online]. Available: https://discoverwater.co.uk/treating-sewage.

[2] Y. Yu, A. Safari, X. Niu, B. Drinkwater and K. V. Horoshenkov, "Acoustic and ultrasonic techniques for defect detection and condition monitoring in water and sewerage pipes: A review," *Applied Acoustics,* vol. 183, no. 0003-682X, p. 108282, 2021.

[3] R. Worley, Y. Yu and S. Anderson, "Acoustic Echo-Localization for Pipe Inspection Robots," in *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2020.

[4] W. Duan, R. Kirby, J. Prisutova and K. V. Horoshenkov, "On the use of power reflection ratio and phase change to determine the geometry of a blockage in a pipe," *Applied Acoustics,* vol. 87, pp. 190-197, 2015.

[5] Y. Yu, A. Krynkin, Z. Li and K. V. Horoshenkov, "Analytical and empirical models for the acoustic dispersion relations in partially filled water pipes," *Applied Acoustics,* vol. 179, no. 0003-682X, p. 108076, 2021.

[6] P. M. Morse and K. U. Ingard, Theoretical acoustics, Princeton university press, 1986.

[7] J. Benesty, J. Chen and Y. Huang, Microphone array signal processing Vol. 1, Springer Science & Business Media, 2008.

[8] M. Ferrante, B. Brunone and S. Meniconi, "Wavelets for the analysis of transient pressure signals

for leak detection," *Journal of hydraulic engineering,* vol. 133, no. 11, pp. 1274-1282., 2007.

[9] J. Owowo and S. O. Oyadiji, "Finite element analysis and experimental measurement of acoustic wave propagation for leakage detection in an air-filled pipe," *International Journal of Structural Integrity,* vol. 8, pp. 452-467, 2017.

[10] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences ,* vol. 365, no. 1851, pp. 303-315, 2007.

[11] G. Kokossalakis, "Acoustic data communication system for in-pipe wireless sensor networks," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA,USA, 2006.

[12] S. L. Brunton and J. N. Kutz., Data-driven science and engineering: Machine learning, dynamical systems, and control., Cambridge University Press, , 2019..

[13] Z. Zhang, Y. Xu, J. Yang, X. Li and D. Zhang, "A survey of sparse representation: algorithms and applications," *IEEE access,* vol. 3, pp. 490-530, 2015.

[14] M. A. T. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *IEEE International Conference on Image Processing 2005. Vol. 2. IEEE*, 2005.

[15] S. Thrun, W. Burgard and F. Dieter, Probabilistic Robotics, 2006.

[16] A. V. Oppenheim, R. W. Schafer and J. R. Buck., Discrete-Time Signal Processing. 2nd Edition, Pearson Education India, 1999.

[17] M. T. B. Ali, K. Horoshenkov and S. J. Tait, "Rapid detection of sewer defects and blockages using acoustic-based instrumentation.," *Water science and technology,* vol. 64, no. 8, pp. 1700-1707, 2011.

[18] S. J. Wright, R. D. Nowak and M. A. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing,* vol. 57(7), pp. 2479-2493, 2009.

1

2