



Article

Leak Detection and Location of Water Pipes Using Vibration Sensors and Modified ML Prefilter

Jihoon Choi 1,* D, Joonho Shin 2, Choonggeun Song 1, Suyong Han 1 and Doo Il Park 2

- School of Electronics and Information Engineering, Korea Aerospace University, Goyang-City, Gyeonggi-do 10540, Korea; choonggeunsong@gmail.com (C.S.); hhssyy444@gmail.com (S.H.)
- Energy Solution/Business Team, LG CNS Co., Ltd., Yeongdeungpo-gu, Seoul 07320, Korea; joonho@lgcns.com (J.S.); dipark@lgcns.com (D.I.P.)
- * Correspondence: jihoon@kau.ac.kr; Tel.: +82-2-300-0417

Received:28 July 2017; Accepted: 11 September 2017; Published: 13 September 2017

Abstract: This paper proposes a new leak detection and location method based on vibration sensors and generalised cross-correlation techniques. Considering the estimation errors of the power spectral densities (PSDs) and the cross-spectral density (CSD), the proposed method employs a modified maximum-likelihood (ML) prefilter with a regularisation factor. We derive a theoretical variance of the time difference estimation error through summation in the discrete-frequency domain, and find the optimal regularisation factor that minimises the theoretical variance in practical water pipe channels. The proposed method is compared with conventional correlation-based techniques via numerical simulations using a water pipe channel model, and it is shown through field measurement that the proposed modified ML prefilter outperforms conventional prefilters for the generalised cross-correlation. In addition, we provide a formula to calculate the leak location using the time difference estimate when different types of pipes are connected.

Keywords: leakage detection; vibration sensor; time difference estimation; generalised cross-correlation; maximum-likelihood (ML) prefilter

1. Introduction

This paper deals with the leakage detection and location of water pipes which is one of the main concerns in the water management field. It has been reported that the amount of non-revenue water reaches 15.2–35.1% of drinking water supply in United States, 6.5–24.6% in Europe, and 4.3–27.0% in Korea [1–3]. These losses can be classified into unbilled public usage, apparent losses including unauthorized consumption and metering inaccuracies, and real losses through overflows at storage tanks and burst leaks in distribution pipelines caused by bad connections, pipe corrosion, and physical damages. It is especially very important to detect the burst leaks in buried water pipes in order to reduce water production costs as well as to protect public safety [1,4,5].

Most of the water pipes are buried underground, making it difficult to find the location of leaks. For this reason, water leakage has usually been detected when water flows out of the ground due to massive leaks in pipes. As a preemptive detection method, an acoustic-based approach is used that an experienced expert scans a suspected area by listening to leak sounds, yet this scheme is extremely time-consuming and the accuracy highly depends on the detection skills of the personnel [6,7]. As a more systematic approach, a supervisory control and data acquisition system (SCADA) has been considered to periodically measure the flow rate or water pressure using permanently installed sensors [8,9]. This technique has evolved into model-based leak detection that compares the measured data taken on the water network with the predicted values from flow models [10,11], and a model falsification method is introduced for efficiently detecting leaks in large-scale water distribution networks [12,13]. These methods identify the areas with significant water leakage, so additional

Sensors **2017**, *17*, 2104 2 of 17

techniques are required to find the exact position of leaks [8,9,12]. In addition, ground penetration radar can be used to locate leaks in water pipes by inspecting the ground surface and detecting voids in the soil caused by leaking water; however, the use of this approach is limited to specific areas because of high costs and high sensitivity to underground metal objects [4,14]. In [15,16], a non-invasive detection method based on the time domain reflectometry (TDR) has been investigated. This approach finds the position of leaks using the change of the reflection coefficient near leakage points. Because the measured reflection coefficient is sensitive to the temperature and material type around the pipe, robust leak detection is not easy.

The most popular approach to water leak detection is to use acoustic/vibration sensors or pressure transducers attached to the surface of a pipe [5,17-29]. When two sensors are used, the leakage location can be determined by estimating the time difference through correlation of the receive signals. The performance of various fast Fourier transform (FFT)-based algorithms is compared in terms of the time difference estimation accuracy in [18–21]; the time difference estimation schemes based on spectral transform are further enhanced by the short-time Fourier transform (STFT) [22] and the wavelet transform [23]; and the correlation-based leak detector has been verified via hardware implementation [24]. In addition, a leak detector is proposed to estimate the time difference using an entropy algorithm [25]. When a single sensor is used, the leakage can be detected using the fact that the change of the frequency spectrum is highly dependent on the change of leakage volume [26]; however, it is not easy to employ this method in real environments due to the sensitivity to the change of frequency response. In [27–29], leak detection schemes with single sensor have been investigated based on the wavelet transform techniques; however, it is difficult to adopt in practical environments because they require reference vibration signals without leaks or the detection performance is vulnerable to the system noise. Additionally, other authors are investigating the use of the filter diagonalization method [30], the radio frequency identification (RFID)-based leakage monitoring system [31], and magnetic flux leakage detection [32]. However, because of the excessive complexity, the difficulty of installation and management, and the technical immaturity, these techniques are rarely used in specific areas.

In this paper, we focus on the time difference estimation method using two vibration sensors attached to water pipes. When the power spectral densities (PSDs) and the cross-spectral density (CSD) of the receive signals are known, the optimal time difference estimator is given by the generalised cross-correlation (GCC) with the maximum likelihood (ML) prefilter derived in [18]. Practically, the PSDs and the CSD include certain estimation errors, and thus the ML prefilter may perform worse than the smoothed coherence transform (SCOT) and Wiener processors [21]. In order to improve the performance of the ML prefilter with erroneous PSDs and CSD, we propose a modified ML method that exploits a regularisation factor to mitigate the estimation errors of the PSDs and CSD. We derive a theoretical variance of the time difference estimation error in the discrete-frequency domain through modification of the variance model in the continuous frequency domain [18], and then find the optimal regularisation factor that minimises the theoretical variance in practical water pipe channels. The time difference estimation accuracy of the proposed method is assessed through computer simulations under the channel similar to an actual water pipe. By implementing the vibration sensor in hardware and the proposed detection method in software, respectively, we carry out field measurement in buried water pipes deployed for household use to verify the performance of the proposed leak detector. Numerical simulations and field tests show that the proposed modified ML method performs better than existing correlation-based techniques, in terms of minimising the root mean square deviation (RMSD) of the time difference estimation error and maximising the peak-to-average ratio (PAR) of the cross-correlation function.

This paper is organized as follows. In Section 2, we derive a formula to find the leak location using a time difference between two vibration sensors when two different kinds of pipes are connected. In Section 3, conventional time difference estimation methods are introduced with the signal model, and the modified ML prefilter is proposed for GCC based on the theoretical variance of time difference

Sensors 2017, 17, 2104 3 of 17

estimation error. The performance of the proposed prefilter is compared with those of existing techniques through computer simulations and field measurement in Sections 4 and 5, respectively. Finally, the conclusions are provided in Section 6.

2. Locating Leaks Using Time Difference

When water is leaked in pipes, internal fluid flows out to generate vibration, and it propagates in both directions from the leak point. As shown in Figure 1, vibration sensors are attached to both ends of the pipes including of a leak point, and thus a difference occurs in the vibration wave reception time due to the difference in distance between the leak point and the sensor. Considering the practical pipeline structure, the material and size of a pipe may be different on both sides of a joint. Define the time difference as

$$\Delta t = t_1 - t_2,\tag{1}$$

where t_1 and t_2 mean the time for the vibration wave to reach sensors 1 and 2 from the leak point, respectively. The time difference Δt can be estimated by correlating the receive signals of sensors 1 and 2. For the time being, it is assumed that Δt is known, and the estimation of Δt is considered in the following section. Suppose that the pipe 1 has a leak as shown in Figure 1a, i.e., $0 \le x \le L_1$. Then, t_1 and t_2 are given by

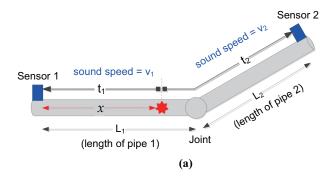
$$t_1 = \frac{x}{v_1}, \tag{2}$$

$$t_1 = \frac{x}{v_1}, (2)$$

$$t_2 = \frac{L_1 - x}{v_1} + \frac{L_2}{v_2}, (3)$$

where x is the distance between the sensor 1 and the leak point; and v_1 and v_2 denote the sound speed in pipes 1 and 2, respectively. From Equations (2) and (3), the time difference is expressed as

$$\Delta t = \frac{2x - L_1}{v_1} - \frac{L_2}{v_2}. (4)$$



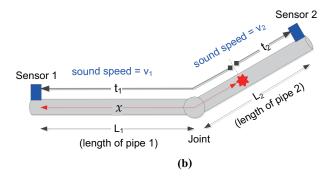


Figure 1. Leak location estimation using the time difference of signals received at sensors, when two different pipes are connected: (a) when pipe 1 has a leak; (b) when pipe 2 has a leak.

Sensors 2017, 17, 2104 4 of 17

By rearranging Equation (4) with respect to x, we have

$$x = \frac{1}{2}(L_1 + L_2 \frac{v_1}{v_2} + v_1 \Delta t),\tag{5}$$

where $\Delta t \leq \frac{L_1}{v_1} - \frac{L_2}{v_2}$. Similarly, when pipe 2 has a leak as shown in Figure 1b, the time difference is expressed as

$$\Delta t = \frac{L_1}{v_1} + \frac{2x - 2L_1 - L_2}{v_2},\tag{6}$$

and the distance between leak point and the sensor 1 is given by

$$x = L_1 + \frac{1}{2}(L_2 + v_2\Delta t - L_1\frac{v_2}{v_1}),\tag{7}$$

where $\Delta t > \frac{L_1}{v_1} - \frac{L_2}{v_2}$. In summary, the leak location from the sensor 1 can be computed as

$$x = \begin{cases} \frac{1}{2}(L_1 + L_2 \frac{v_1}{v_2}) + \frac{v_1}{2} \Delta t, & \text{for } \Delta t_{min} \le \Delta t \le \frac{L_1}{v_1} - \frac{L_2}{v_2}, \\ L_1(1 - \frac{v_2}{2v_1}) + \frac{1}{2}L_2 + \frac{v_2}{2} \Delta t, & \text{for } \frac{L_1}{v_1} - \frac{L_2}{v_2} < \Delta t \le \Delta t_{max}, \end{cases}$$
(8)

where $\Delta t_{max} = \frac{L_1}{v_1} + \frac{L_2}{v_2}$ and $\Delta t_{min} = -\Delta t_{max}$. The speed of sound in a water-filled pipe depends on the properties of a pipe wall that the vibration wave is traveling through. In [33], the velocity of a fluid-dominated wave is expressed as

$$v = v_f \left(1 + \frac{(2B_f/a)}{(Eh/a^2) - \rho h\omega^2} \right)^{-\frac{1}{2}},\tag{9}$$

where B_f is the bulk modulus of the contained medium; a, h, and E are the inner radius, the thickness, and the elastic modulus of the pipe, respectively; ρ is the density of the surrounding medium; ω is the angular velocity; and v_f is the free-field internal fluid velocity [34], given by

$$v_f = \sqrt{\frac{K_s}{\rho}},\tag{10}$$

where K_s is a stiffness coefficient of the material. We use the database of the sound speed created for various types of pipes by theoretical calculations and experimental corrections provided by [35], and Table 1 presents a part of the database. Practically, the sound speed varies considerably according to the geometry and material type of the pipe. If a pipe is composed of mixed media, it is difficult to calculate the sound speed. Thus, we only consider a single material pipe.

Table 1. Sound ve	elocity accordin	g to the pij	pe material and	diameter [35].
--------------------------	------------------	--------------	-----------------	----------------

Material	Diameter (mm)	Velocity (m/s)	
	40	565	
Polyvinyl Chloride (PVC)	80	540	
	150	530	
	150	1220	
Cast-Iron	250	1160	
	350	1120	
	25	1375	
	40	1350	
Steel	60	1330	
	90	1286	
	150	1200	
	250	1150	

Sensors **2017**, *17*, 2104 5 of 17

3. Time Difference Estimation

3.1. Conventional Time Difference Estimation Methods

In Figure 1, define $x_1(n)$ and $x_2(n)$ as the sampled receive signals of sensors 1 and 2 at time $t = nT_s$, respectively, where T_s is the sampling period. When there is a leak between sensors 1 and 2, $x_1(n)$ and $x_2(n)$ are given by

$$x_1(n) = h_1(n) \otimes s(n) + w_1(n),$$
 (11)

$$x_2(n) = h_2(n) \otimes s(n+d) + w_2(n),$$
 (12)

where s(n) is the vibration wave generated from the leak point; d is the time difference in samples between received signals at sensors 1 and 2 that satisfies $\Delta t = dT_s$; and \otimes means the discrete-time convolution operator. For $i \in \{1,2\}$, $h_i(n)$ is the discrete-time channel impulse response between the leak point and the sensor i, and $w_i(n)$ is a stationary random process representing the background noise at the sensor i. It is assumed that $x_1(n)$ and $x_2(n)$ are two stationary random signals uncorrelated with $w_i(n)$, and also assumed that $w_1(n)$ and $w_2(n)$ are uncorrelated. The time difference can be estimated using the cross-correlation of $x_1(n)$ and $x_2(n)$ defined as

$$R_{x_1x_2}(\tau) = E[x_1(n)x_2^*(n-\tau)],\tag{13}$$

where τ is the time delay of $x_2(n)$, $E[\cdot]$ denotes the expectation operator, and $(\cdot)^*$ means the complex conjugate.

The cross-correlation function $r_{x_1x_2}(\tau)$ can be computed in the time domain by directly correlating $x_1(n)$ and $x_2(n)$ because $R_{x_1x_2}(\tau)$ has the peak value at $\tau=d$. This approach, however, is very sensitive to the non-leak vibration noises, and requires high computational load when the sample size is large. For these reasons, the time difference estimation method in the frequency domain is widely used based on the relationship between the cross-correlation function and the CSD [18–24]. Let us define DFT $_N[x(n)]$ as the N-point discrete Fourier transform (DFT) of a sequence x(n). Then, the CSD for (13) is written as

$$S_{x_1x_2}(k) = E[X_1(k)X_2^*(k)]$$

= $H_1(k)H_2^*(k)S_s(k)e^{-j\frac{2\pi kd}{N}},$ (14)

where $X_i(k) = \mathrm{DFT}_N[x_i(n)]$ and $H_i(k) = \mathrm{DFT}_N[h_i(n)]$ for $i \in \{1,2\}$; $S_s(k)$ is the PSD of s(k); and $-\frac{N}{2} \le k \le \frac{N}{2} - 1$. By computing $S_{x_1x_2}(k)$ and taking its inverse DFT (IDFT), the cross-correlation function can be simply computed when the sample size is large.

To further enhance the frequency domain estimation method by sharpening the peak in the cross-correlation function, prefiltering is performed before the IDFT. In this case, the cross-correlation function is computed using the GCC method [18], defined as

$$R_{x_1x_2}^g(\tau) = IDFT_N[\Psi(k)S_{x_1x_2}(k)], \tag{15}$$

where $\mathrm{IDFT}_N[x(k)]$ means the N-point IDFT of a sequence x(k) and $\Psi(k)$ is a frequency domain prefilter. Figure 2 describes the procedure to compute the cross-correlation function in the GCC method, and Table 2 presents the prefilters for the GCC proposed in [18–20], where $S_{w_i}(k)$ and $S_{x_i}(k)$ are the PSDs of $w_i(n)$ and $x_i(n)$ for $i \in \{1,2\}$, respectively; and $\gamma(k)$ is the coherence function between $x_1(n)$ and $x_2(n)$ defined as

$$\gamma(k) = \frac{S_{x_1 x_2}(k)}{\sqrt{S_{x_1}(k)S_{x_2}(k)}}.$$
(16)

Sensors **2017**, *17*, 2104 6 of 17

In the Wiener and ML prefilters, the background noises $w_1(n)$ and $w_2(n)$ are assumed to be zero-mean Gaussian random variables with variances σ_1^2 and σ_2^2 , respectively, to make the signal models (11) and (12) mathematically tractable. As pointed out in [12,13], $E[w_i(n)]$ is not zero due to the sensor bias, and the modeling error included in $w_i(n)$ may not be expressed as a Gaussian distribution. The effect of the bias in $w_i(n)$ can be removed by sensor calibration or subtracting the mean value from $x_i(n)$. In addition, for a large N, $W_i(k) = \mathrm{DFT}_N[w_i(k)]$ tends to be Gaussian by the central limit theorem, even if $w_i(n)$ is not a Gaussian random variable. Hence, the assumption that $w_1(n)$ and $w_2(n)$ are Gaussian is not a strong requirement ([36], Ch.3A).

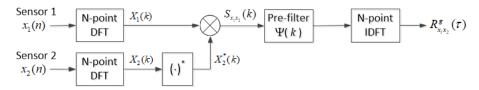


Figure 2. Block diagram to compute $R_{x_1x_2}^g(\tau)$ using the generalised cross-correlation method.

Table 2. Conventiona	l prefilters for the	generalised cross	-correlation method	[18-20].
----------------------	----------------------	-------------------	---------------------	----------

Prefilter Type	Weight Function
Roth	$\frac{1}{S_{x_1}(k)}$
Smoothed coherence transform (SCOT)	$\frac{1}{\sqrt{S_{x_1}(k)S_{x_2}(k)}}$
Phase transform (PHAT)	$\frac{1}{ S_{x_1x_2}(k) }$
Eckart	$\frac{ S_{x_1x_2}(k) }{S_{w_1}(k)S_{w_2}(k)}$
Wiener	$ \gamma(k) ^2$
Maximum-likelihood (ML)	$\frac{ \gamma(k) ^2}{1 - \gamma(k) ^2} \frac{1}{ S_{x_1 x_2}(k) }$

The Roth prefilter proposed by Peter R. Roth normalises the CSD by $S_{x_1}(k)$, while the smoothed coherence transform (SCOT) prefilter normalises the CSD by $\sqrt{S_{x_1}(k)S_{x_2}(k)}$. The Roth and SCOT prefilters can be used only in the high signal-to-noise ratio (SNR) region because the peak of the cross-correlation function is spread out after prefiltering. To avoid the peak spreading, the phase transform (PHAT) normalises the CSD by $|S_{x_1x_2}(k)|$ so that the prefiltered CSD has the unit amplitude for all k. However, the peak estimation performance is degraded due to the noise enhancement in the frequency band with low SNR. The Eckart prefilter maximises the ratio of the mean correlator output power to the noise variance of the correlator output, and thus it suppresses the frequency band with low SNR. This method is difficult to use in a practical system because it is not easy to estimate the noise PSDs $S_{w_1}(k)$ and $S_{w_2}(k)$ from the receive signals with leaks [18,20,21]. The ML prefilter weights the CSD according to the SNR in each frequency band to minimise the variance of the time difference estimate, yet it may critically overweight the CSD in the frequency band where $|\gamma_{x_1x_2}(k)|^2$ is close to one [18]. The Wiener prefilter multiplies the CSD by the weight $|\gamma_{x_1x_2}(k)|^2$, so that it mitigates the overemphasising problem at certain frequencies [19].

3.2. Proposed Time Difference Estimation Scheme Using the Modified ML Method

For practical implementation of the prefilter, it is required to estimate the PSDs $S_{x_1}(k)$ and $S_{x_2}(k)$ as well as the CSD $S_{x_1x_2}(k)$. Additionally, the Eckart prefilter necessitates the estimation of the noise PSDs $S_{w_1}(k)$ and $S_{w_2}(k)$. Suppose that we have long sequences $x_1(n)$ and $x_2(n)$ measured at sensors 1

Sensors **2017**, *17*, 2104 7 of 17

and 2. Then, the PSDs and CSD of $x_1(n)$ and $x_2(n)$ can be estimated by the non-overlapped Welch's method [37] as follows:

$$\hat{S}_{x_i}(k) = \frac{1}{M} \sum_{m=1}^{M} |X_i^{(m)}(k)|^2, \tag{17}$$

$$\hat{S}_{x_1 x_2}(k) = \frac{1}{M} \sum_{m=1}^{M} X_1^{(m)}(k) \left(X_2^{(m)}(k) \right)^*, \tag{18}$$

where $X_i^{(m)}(k) = \text{DFT}_N[x_i(\frac{mN}{2}), x_i(\frac{mN}{2}+1), \cdots, x_i(\frac{(m+1)N}{2}-1)];$ M is the number of blocks; and $i \in \{1, 2\}$. Using the power spectrum estimates in Equation (18), the ML prefilter is expressed as

$$\Psi(k) = \frac{|\hat{\gamma}(k)|^2}{1 - |\hat{\gamma}(k)|^2} \frac{1}{|\hat{S}_{x_1 x_2}(k)|'}$$
(19)

where $\hat{\gamma}(k)$ is the estimated coherence function, given by

$$\hat{\gamma}(k) = \frac{\hat{S}_{x_1 x_2}(k)}{\sqrt{\hat{S}_{x_1}(k)\hat{S}_{x_2}(k)}}.$$
(20)

Note that $0 \le |\hat{\gamma}(k)| \le 1$. When a GCC technique is used to estimate the time difference, the error variance has been derived using the integration of a weighted CSD function in the frequency domain [18]. In the discrete-frequency domain, the error variance for an arbitrary prefilter $\Psi(k)$ is obtained through the summation of a weighted discrete CSD function as follows:

$$\operatorname{var}(\hat{d}) = \frac{N^2}{(2\pi)^2} \frac{\sum_{k=-N/2}^{N/2-1} k^2 \Psi^2(k) [1 - |\gamma(k)|^2] S_{x_1}(k) S_{x_2}(k)}{\left[\sum_{k=-N/2}^{N/2-1} k^2 \Psi(k) |\hat{S}_{x_1 x_2}(k)|\right]^2}.$$
 (21)

To derive the error variance of the ML prefilter using the estimated PSDs and CSD, we substitute Equation (19) into Equation (21). In addition, by approximating that $|S_{x_1x_2}(k)|\cong |\hat{S}_{x_1x_2}(k)|$, we get

$$\operatorname{var}(\hat{d}) = \frac{N^{2}}{(2\pi)^{2}} \frac{\sum_{k=-N/2}^{N/2-1} k^{2} \hat{\beta}_{0}(k) [1 - |\gamma(k)|^{2}] \frac{S_{x_{1}}(k) S_{x_{2}}(k)}{|\hat{S}_{x_{1}x_{2}}(k)|^{2}}}{\left[\sum_{k=-N/2}^{N/2-1} k^{2} \hat{\beta}_{0}^{2}(k)\right]^{2}}$$

$$\cong \frac{N^{2}}{(2\pi)^{2}} \frac{\sum_{k=-N/2}^{N/2-1} k^{2} \hat{\beta}_{0}^{2}(k) \frac{1}{\beta_{0}(k)}}{\left[\sum_{k=-N/2}^{N/2-1} k^{2} \hat{\beta}_{0}(k)\right]^{2}},$$
(22)

where $\beta_0(k) = |\gamma(k)|^2/(1-|\gamma(k)|^2)$, and $\hat{\beta}_0(k)$ is the estimate of $\beta_0(k)$ given by

$$\hat{\beta}_0(k) = |\hat{\gamma}(k)|^2 / (1 - |\hat{\gamma}(k)|^2). \tag{23}$$

Sensors 2017, 17, 2104 8 of 17

It can be easily shown that $var(\hat{d})$ is minimised when $\hat{\beta}_0(k) = \beta_0(k)$ for all k. Practically, $\hat{\beta}_0(k)$ is highly dependent on the estimation error of $\gamma(k)$. Especially, when $|\hat{\gamma}(k)|$ is close to one, $\hat{\beta}_0(k)$ is very sensitive to the variation of $\hat{\gamma}(k)$, thereby resulting in a rapid increase of $var(\hat{d})$.

In an attempt to mitigate the influence of coherence estimation error, we propose a new modified ML prefilter that employs a regularisation factor when computing $\hat{\beta}_0(k)$, and thus it is more robust to the coherence estimation error than the conventional ML prefilter. The proposed modified ML prefilter is expressed as

$$\Psi(k) = \hat{\beta}(k) \frac{1}{|\hat{S}_{x_1 x_2}(k)|}, \tag{24}$$

where $\hat{\beta}(k)$ is defined as

$$\hat{\beta}(k) = \frac{|\hat{\gamma}(k)|^2}{1 - |\hat{\gamma}(k)|^2 + \alpha}.$$
(25)

Here, $0 < \alpha \le 1$ is the regularisation factor. Since the denominator of Equation (25) is equal to or greater than α , the estimation error included in $\hat{\gamma}(k)$ gives less impact on the change of $\hat{\beta}(k)$ compared to the conventional ML prefilter. The proposed modified ML prefilter (24) is equal to the ML prefilter when $\alpha = 0$, while it is equal to the Wiener prefilter when $\alpha = |\hat{\gamma}(k)|^2$. By replacing $\hat{\beta}_0(k)$ in Equation (22) with $\hat{\beta}(k)$ in Equation (25), the error variance for the modified ML method is denoted as

$$\operatorname{var}(\hat{d}) = \frac{N^2}{(2\pi)^2} \frac{\sum_{k=-N/2}^{N/2-1} k^2 \hat{\beta}^2(k) \frac{1}{\beta_0(k)}}{\left[\sum_{k=-N/2}^{N/2-1} k^2 \hat{\beta}(k)\right]^2}.$$
 (26)

As the regularization factor α increases, $\hat{\beta}(k)$ becomes more robust to the estimation error of $\gamma(k)$; however, the bias of $\hat{\beta}(k)$ increases. In addition, the optimal α minimising $\mathrm{var}(\hat{d})$ is varied according to SNR values of receive signals and statistical characteristics of $\gamma(k)$. Therefore, we carefully design α considering the variance of time difference error in numerical simulations and field measurements. The specific design method for α is described in Sections 4 and 5.

4. Simulation Results

4.1. Channel Model and Optimisation of Regularisation Factor

To perform computer simulations, we construct a channel model for vibration signals in buried water pipes based on the field measurement data. Details of field measurements are explained in Section 5. We define a channel model for propagation of vibration signals in a water pipe as below:

$$H_i(z) = \left(\frac{d_i}{d_0}\right)^{-m} \frac{1}{1 + c_1 z^{-K}} \frac{(1 - c_2^2)^{1/2}}{1 - c_2 z^{-1}} + E_i(z),\tag{27}$$

where $d_0=1$ m is the reference distance; d_i is the distance between the leak point and the sensor i; m is the path loss exponent; $-1 \le c_1 \le 1$ and K are the coefficient and order of a comb filter; $0 < c_2 < 1$ is the coefficient of a first-order infinite impulse response (IIR) filter; and $E_i(z)$ is an error term describing the channel variation. It is assumed that, for various ω , $E_i(e^{j\omega})$ is an independent and identically distributed (i.i.d.) Gaussian random variable with zero mean and variance σ_h^2 and that $E_1(e^{j\omega})$ and $E_2(e^{j\omega})$ are independent. In Equation (27), $1/(1+c_1z^{-K})$ is the inverse comb filter to represent a sharp increase in the frequency response at harmonics of the natural frequency, and the first-order IIR filter describes the lowpass characteristics of vibration signals.

Figure 3 compares the coherence function of the signals generated by the channel model (27) and that of the measured signals in the field test, when N=4096 and M=30. For the channel model, we used $d_1=20$ m, $d_2=30$ m, m=2, $c_1=0.7$, K=69, $C_2=0.7$, and $C_0=0.02$. In the field

Sensors **2017**, *17*, 2104 9 of 17

measurement, the coherence is obtained by averaging three trials of measurement data for Location 3 and sensors 1 and 3 (L3S13). The two coherence values are very similar except that the normalised frequency is around 0.5, and this shows the validity of the proposed channel model.

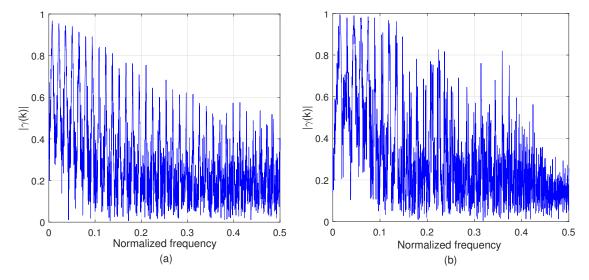


Figure 3. Coherence $\gamma(k)$ when N=4096, and M=30: (a) signals generated by the channel model (27); (b) measured signals in the field test for Location 3 and sensors 1 and 3 (L3S13).

Using the channel model (27), we find the optimal regularisation factor α for the proposed modified ML prefilter. In the simulation, the coherence $\gamma(k)$ is generated using the same parameters as in Figure 3, and the coherence estimate $\hat{\gamma}(k)$ is denoted as

$$\hat{\gamma}(k) = \gamma(k) + \gamma_e(k),\tag{28}$$

where $\gamma_e(k)$ is the estimation error of $\gamma(k)$ that is assumed to be a zero-mean Gaussian random variable with variance σ_e^2 . Figure 4 presents the normalised variance of time difference estimation error, which is defined as

$$\operatorname{var}_{N}(\hat{d};\alpha) = \frac{\operatorname{var}(\hat{d};\alpha)}{\operatorname{var}(\hat{d};\alpha=0)},\tag{29}$$

where $\mathrm{var}(\hat{d};\alpha)$ is computed by Equations (25) and (26). The normalised error variance monotonically decreases as the regularisation factor α increases to 0.5, whereas it grows with increment of α when $\alpha>0.6$. Thus, the normalised error variance is minimised when $0.5 \leq \alpha \leq 0.6$, and the optimal α is slightly different depending on σ_e^2 . As σ_e^2 decreases, the minimum of normalised error variance decreases, i.e., the performance gain of the modified ML prefilter increases compared to the conventional ML prefilter with $\alpha=0$.

As mentioned in [12,13], $\gamma_e(k)$ may have a non-Gaussian distribution with a bias in a practical system. However, to the authors' best knowledge, the distribution of $\gamma_e(k)$ in practical water pipes are not available in the literature, and thus $\gamma_e(k)$ is assumed to be zero-mean Gaussian in the simulation. Since $\operatorname{var}(\hat{d})$ is a function of $\hat{\gamma}(k)$, the normalised error variance curve in Figure 4 is varied according to the distribution of $\gamma_e(k)$ and the optimal α also can be changed. For more practical design, we compute the normalised error variance curve and the corresponding optimal α using $\hat{\gamma}(k)$ measured at real test sites in Section 5.2.

Sensors 2017, 17, 2104 10 of 17

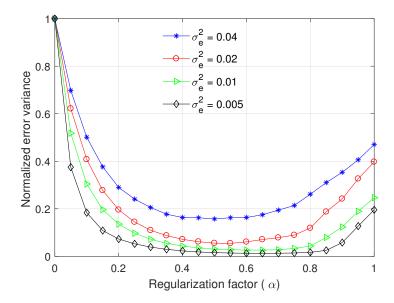


Figure 4. Normalised variance of time difference error as a function of α , when N = 4096, and M = 30.

4.2. Performance Comparison

In this subsection, we compare the performance of the proposed modified ML method with those of existing prefilters. As mentioned in [18,21], the PHAT prefilter performs better than the Roth and SCOT prefilters with the peak spreading problem. In addition, as explained in Section 3, the Eckart prefilter requires the noise PSDs $S_{w_1}(k)$ and $S_{w_2}(k)$, which are difficult to estimate in a practical system. For these reasons, we consider the PHAT, Wiener, and ML prefilters for comparison with the proposed method. To compare the performance of various prefilters, we utilize the RMSD of the time difference estimate given by

$$RMSD(\Delta \hat{t}) = T_s \sqrt{E[|\hat{d} - d|^2]}.$$
(30)

Note that $RMSD(\Delta \hat{t}) = T_s \sqrt{var(\hat{d})}$ when \hat{d} is unbiased. In addition, we employ the PAR to describe the relative magnitude of the maximum peak compared to the average of the cross-correlation function. Specifically, when the maximum peak appears at $\tau = \hat{d}$, the PAR is defined as

$$PAR(\hat{d}) = \frac{|R_{x_1 x_2}^{g}(\hat{d})|^2}{\frac{1}{N} \sum_{\tau = -N/2}^{N/2 - 1} |R_{x_1 x_2}^{g}(\tau)|^2},$$
(31)

where the cross-correlation function $R_{x_1x_2}^g(\tau)$ is given by Equation (15).

Figures 5 and 6 show the RMSD and the PAR for various prefilters, respectively, when the channel model (27) is used, N=4096, and M=30. α was set to 0.5 for the proposed modified ML prefilter; σ_h^2 was determined by the SNR at sensor 1; d was randomly generated in the range between -200 and 200; and other parameters were the same as in Figure 4. All curves were obtained by averaging the simulation results over 10^4 channel realizations. The RMSD decreases with increment of SNR, and the proposed method outperforms the conventional PHAT, ML, and Wiener prefilters in the whole SNR region. The RMSD value of the PHAT prefilter approaches those of the ML and Wiener prefilters as the SNR increases, because the PHAT prefilter has good performance in high SNR region [18,21]. In Figure 6, the proposed modified ML method presents the highest PAR among all prefilters for entire SNR region, i.e., the proposed prefilter can detect the leakage point more reliably than other techniques. The PAR of the ML prefilter is saturated when SNR ≥ 4 dB because the estimation error of $\gamma(k)$ influences both the numerator and denominator of Equation (23).

Sensors 2017, 17, 2104 11 of 17

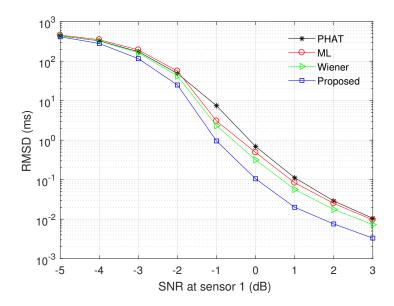


Figure 5. Root mean square deviation (RMSD) of time difference error for various prefilters, when $\alpha = 0.5$, N = 4096, and M = 30.

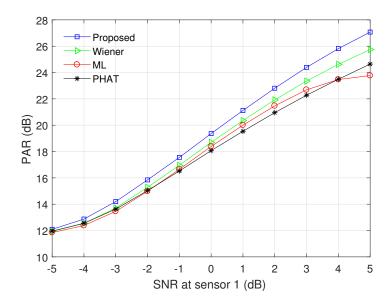


Figure 6. Peak-to-average ratio (PAR) of the time domain cross-correlation function for various prefilters, when $\alpha = 0.5$, N = 4096, and M = 30.

5. Field Measurement

5.1. System Setup for Field Testing

Figure 7 shows the leak detection system for field measurements. We developed a web-based leakage detection system including our own sensor composed of a vibration transducer, a microprocessor, and a communication module. Once the control node sends the *start logging* command to the sensors via the network node, the sensors record the vibration signals for 30 s and transmit the sampled data to the server through the wireless communication link and the commercial internet protocol (IP) network. Prior to data logging, the network node broadcasts a sync message, and the local clock of the sensor is synchronized using the message to avoid the time difference estimation error caused by the time offset between sensors. The measured data stored in the server can be extracted for leakage detection at the control node. At the sensor, the sampling frequency was 4 kHz

Sensors **2017**, *17*, 2104

($T_s = 0.25$ ms) and each sample is represented as 12 bits. The wireless link between the sensor and the network node supports 40 kbps of data rate using binary phase shift-keying (BPSK) modulation.

Field measurements were carried out in Seogwipo City, Korea, in order to verify the leak detection performance of the proposed method. To determine specific test areas, we utilized the leak rate information in [3] and the pipeline database provided by the local government. We have selected five locations to perform field measurements as shown in Figure 8, and specific pipe layouts in these locations are given by Figure 9. The leaks in water pipes occur mostly at the joint to connect pipes and the branch to distribute water to small pipes, and these leak points can be identified by checking manholes and the entrance of household pipes. Thus, we have found the actual leak point described as red spots in Figure 9 through the inspection of possible leak points as well as the leak detection based on the time difference estimation. The sound velocity in Figure 9 was determined by querying the database provided by [35] using the information about the material and diameter of the water supply pipe. Since two adjacent sensors are selected to log vibration signals for each field trial, there are nine kinds of test cases as shown in Table 3. For notational convenience, we use the test case names in Table 3 in Sections 4 and 5.

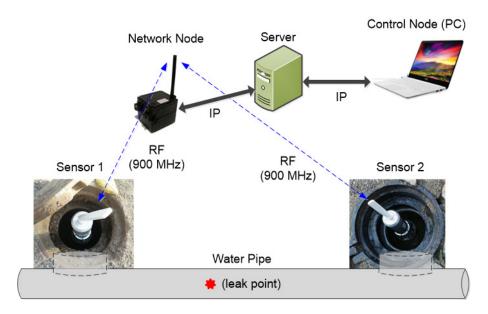


Figure 7. Leak detection system for field measurements.

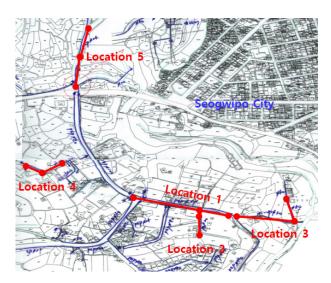


Figure 8. Locations for field measurements.

Sensors **2017**, *17*, 2104

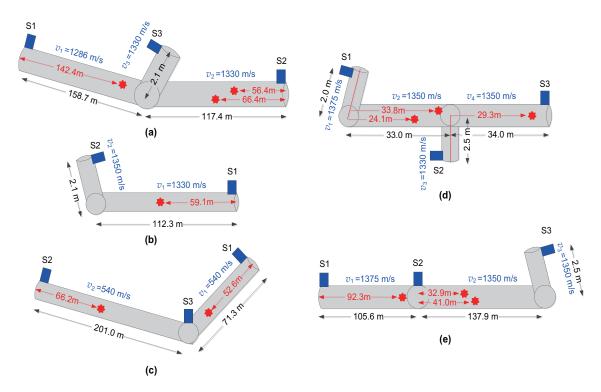


Figure 9. Pipe layout used in field measurements where S1, S2, and S3 means Sensors 1, 2, and 3; and red spots denote leak points: (a) Location 1; (b) Location 2; (c) Location 3; (d) Location 4; (e) Location 5.

Table 3. Test cases in the field measurement.

Test Case	L1S13	L1S23	L2S12	L3S13	L3S23	L4S12	L4S23	L5S12	L5S23
Location	1	1	2	3	3	4	4	5	5
Sensor Pair	S1, S3	S2, S3	S1, S2	S1, S3	S2, S3	S1, S2	S2, S3	S1, S2	S2, S3

5.2. Performance Evaluation

In a practical system, the real coherence $\gamma(k)$ is not available. Instead, $\gamma(k)$ is estimated using the received signals $x_1(n)$ and $x_2(n)$ through multiple trials of measurement. Using $\gamma(k)$ measured at five test locations, we compare the normalised variance of time difference error as a function of α as shown in Figure 10. The coherence function $\gamma(k)$ was obtained by averaging three separate measurement data for each test case. We used N=4096 and M=25, and σ_e^2 was set to 0.02. For varying α , the normalised error variance presents huge variations depending on a test place because of the change of $\gamma(k)$ and some interference by the external environment; however, the normalised error variance changes slightly when $0.30 \le \alpha \le 0.55$ having the minimum value between 0.35 and 0.60 as shown in Table 4. Based on the analysis in Figure 10, α was set to 0.45 for the proposed prefilter.

Table 4. Optimal regularisation factor minimising the time difference error variance in field measurements.

Test Case	L1S23	L2S12	L3S23	L4S12	L5S12
Optimal α	0.35	0.40	0.60	0.35	0.45

Sensors 2017, 17, 2104 14 of 17

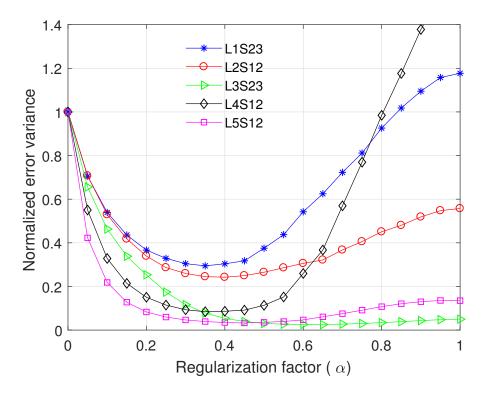


Figure 10. Normalised variance of time difference error as a function of α , when measured coherence values were used, N=4096, M=25, and $\sigma_e^2=0.02$.

As mentioned in Section 5.1, the field measurement was carried out for nine test cases in five locations. The field measurement was performed three times for each test case. We used N=4096and M = 25 for estimating the PSDs $S_{x_1}(k)$ and $S_{x_2}(k)$, and the CSD $S_{x_1x_2}(k)$; and α was set to 0.45 according to the analysis in Section 4.1. Based on the measurement data, Figures 11 and 12 present the RMSD and the PAR of various prefilters, respectively. The proposed modified ML prefilter achieves the smallest RMSD value in all test cases except L1S13 in which the PHAT prefilter has slightly less RMSD than the proposed method. Note that, in the test case L1S13, the RMSD values of all prefilters are less than 0.6 ms, i.e., the proposed method achieves proper estimation accuracy for finding the leak location. The RMSD of the proposed method is less than 1.1 ms for entire test cases. In Figure 12, the proposed prefilter obtains the highest PAR in all test cases except L2S12 in which the PAR value of the PHAT prefilter is 0.7 dB greater than that of the proposed scheme. The PHAT and Wiener prefilters have similar PAR values; however, the RMSD of the PHAT prefilter is higher than that of the Wiener method as shown in Figure 11. In addition, while the ML and Wiener prefilters have similar RMSD values except L5S23, the ML method presents much lower PAR values than the Wiener prefilter. Thus, the Wiener prefilter is better than the PHAT and ML methods. Overall, the proposed method outperforms the conventional prefilters because it achieves the minimum RMSD value among all prefilters while maximising the PAR. In other words, when the distance between sensors are identical, the proposed prefilter has higher leak detection probability than the conventional methods. In addition, under a fixed RMSD requirement, the distance between sensors can be longer in the proposed prefilter than those in the conventional schemes, and thus the sensor deployment cost can be reduced.

Sensors 2017, 17, 2104 15 of 17

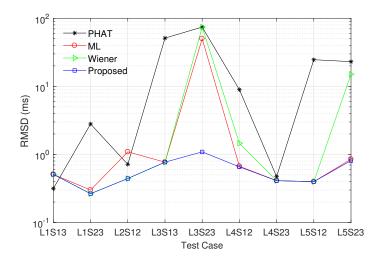


Figure 11. RMSD of various prefilters in field measurements, when $\alpha = 0.45$, N = 4096, and M = 25.

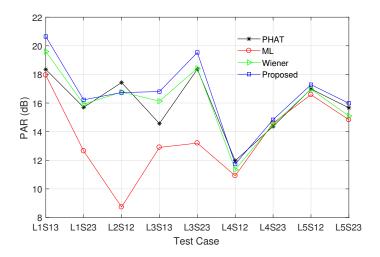


Figure 12. PAR of various prefilters in field measurements, when $\alpha = 0.45$, N = 4096, and M = 25.

6. Conclusions

This paper proposed a new time difference estimation method for leakage detection based on the modified ML prefilter with the regularisation factor. Through analysis and numerical simulations, we have found the optimal regularisation factor and compared the performance of the proposed prefilter with that of conventional prefilters. The proposed leak detection method was verified by field measurements using a practical leak detection system, and it was shown that the proposed method outperforms the conventional techniques in terms of RMSD and PAR. The proposed modified ML prefilter can be applied to correlation-based leak detection systems utilizing the STFT and the wavelet transform. Furthermore, the proposed method can contribute to developing an automatic leakage management solution that collects leakage data, alarms about the risk of leaks, and informs about the specific leak locations.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B4013418) and the R&D project of LG CNS Co., Ltd.

Author Contributions: Jihoon Choi designed the conceptual algorithm and performed computer simulations; Choonggeun Song and Suyong Han evaluated the performance of leak detection methods through computer simulations and analyzed the measurement data; Joonho Shin and Doo II Park were in charge of the field measurement providing guidance for this research; Jihoon Choi and Joonho Shin wrote and revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Sensors 2017, 17, 2104 16 of 17

References

1. United States EPA Office of Water. Control and Mitigation of Drinking Water Losses in Distribution Systems. Available online: https://nepis.epa.gov/ (accessed on 20 July 2017).

- 2. BDEW German Association of Energy and Water Industries. VEWA Survey—Comparison of European Water and Wastewater Prices. Available online: https://www.bdew.de/ (accessed on 20 July 2017).
- 3. Ministry of Environment, Republic of Korea. 2015 Statistics of Waterworks. Available online: http://me.go.kr/(accessed on 30 June 2017).
- 4. Puust, R.; Kapelan, Z.; Savic, D.A.; Koppel, T. A Review of Methods for Leakage Management in Pipe Networks. *Urban Water J.* **2010**, *7*, 25–45, doi:10.1080/15730621003610878.
- 5. Martini, A.; Troncossi, M.; Rivola, A. Automatic Leak Detection in Buried Plastic Pipes of Water Supply Networks by Means of Vibration Measurements. *Shock Vib.* **2015**, 2015, 1–13, doi:10.1155/2015/165304.
- 6. Liston, D.A.; Liston, J.D. Leak Detection Techniques. J. N. Engl. Water Works Assoc. 1992, 106, 103–108.
- 7. Hunaidi, O. Detecting leaks in water-distribution pipes. Constr. Technol. Updates 2000, 40, 1-6.
- 8. Hunaidi, O.; Chu, W.; Wang, A.; Guan, W. Detecting leaks in plastic pipes. *J. Am. Water Works Assoc.* **2000**, 92, 82–94.
- 9. Mounce, S.R.; Boxall, J.B.; Machell, J. Development and Verification of an Online Artificial Intelligence System for Detection of Bursts and Other Abnormal Flows. *J. Water Resour. Plan. Manag.* **2010**, 136, 309–318, doi:10.1061/(ASCE)WR.1943-5452.0000030.
- 10. Furthermoreersen, J.H.; Powell, R.S. Implicit State-Estimation Technique for Water Network Monitoring. *Urban Water* **2000**, 2, 123–130, doi:10.1016/S1462-0758(00)00050-9.
- 11. Poulakis, Z.; Valougeorgis, D.; Papadimitriou, C. Leakage Detection in Water Pipe Networks Using a Bayesian Probabilistic Framework. *Probab. Eng. Mech.* **2003**, *18*, 315–327, doi:10.1016/S0266-8920(03)00045-6.
- 12. Goulet, J.A.; Coutu, S.; Smith, I.F. Model Falsification Diagnosis and Sensor Placement for Leak Detection in Pressurized Pipe Networks. *Adv. Eng. Inf.* **2013**, 27, 261–269, doi:10.1016/j.aei.2013.01.001.
- 13. Moser, G.; Paal, S.G.; Smith, I.F. Performance Comparison of Reduced Models for Leak Detection in Water Distribution Networks. *Adv. Eng. Inf.* **2015**, *29*, 714–726, doi:10.1016/j.aei.2015.07.003.
- 14. Ayala–Cabrera, D.; Herrera, M.; Izquierdo, J.; Ocaña–Levario, S.J.; Pérez–García, R. GPR-Based Water Leak Models in Water Distribution Systems. *Sensors* **2013**, *13*, 15912–15936, doi:10.3390/s131215912.
- 15. Cataldo, A.; Cannazza, G.; Benedetto, E.D.; Giaquinto, N. A New Method for Detecting Leaks in Underground Water Pipelines. *IEEE Sens. J.* 2012, 12, 1660–1667, doi:10.1109/JSEN.2011.2176484.
- 16. Giaquinto, N.; D'Aucelli, G.M.; Benedetto, E.D.; Cannazza, G.; Cataldo, A.; Piuzzi, E.; Masciullo, A. Criteria for Automated Estimation of Time of Flight in TDR Analysis. *IEEE Trans. Instrum. Meas.* **2016**, 65, 1215–1224, doi:10.1109/TIM.2015.2495721.
- 17. Hunaidi, O.; Wang, A.; Bracken, M. Acoustic Methods for Locating Leaks in Municipal Water Pipe Networks. In Proceedings of the International Conference Water Demand Management, Dead Sea, Jordan, 30 May–3 June 2004; pp. 1–14.
- 18. Knapp, C.; Carter, G. The Generalized Correlation Method for Estimation of Time Delay. *IEEE Trans. Acoust. Speech Signal Process.* **1976**, 24, 320–327, doi:10.1109/TASSP.1976.1162830.
- 19. Hero, A.; Schwartz, S. A New Generalized Cross Correlator. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, 33, 38–45, doi:10.1109/TASSP.1985.1164515.
- 20. Carter, G.C. Coherence and Time Delay Estimation. *Proc. IEEE* 1987, 75, 236–255, doi:10.1109/PROC. 1987.13723.
- 21. Gao, Y.; Brennan, M.; Joseph, P. A Comparison of Time Delay Estimators for the Detection of Leak Noise Signals in Plastic Water Distribution Pipes. *J. Sound Vib.* **2006**, 292, 552–570, doi:10.1016/j.jsv.2005.08.014.
- 22. Lay-Ekuakille, A.; Vendramin, G.; Trotta, A.; Vanderbemden, P. STFT-Based Spectral Analysis of Urban Waterworks Leakage Detection. In Proceedings of the XIX IMEKO World Congress Fundamental and Applied Metrology, Lisbon, Portugal, 6–11 September 2009.
- 23. Ge, C.; Yang, H.; Ye, H.; Wang, G. A Fast Leak Locating Method Based on Wavelet Transform. *Tsinghua Sci. Technol.* **2009**, *14*, 551–555, doi:10.1016/S1007-0214(09)70116-6.
- 24. Zhang, L.; Wu, Y.; Guo, L.; Cai, P. Design and Implementation of Leak Acoustic Signal Correlator for Water Pipelines. *Inf. Technol. J.* **2013**, *12*, 2195–2200, doi:10.3923/itj.2013.2195.2200.

Sensors **2017**, *17*, 2104

25. Yang, J.; Wen, Y.; Li, P. Information Processing for Leak Detection on Underground Water Supply Pipelines. In Proceedings of the 3rd International Workshop on Advanced Computational Intelligence, Suzhou, China, 25–27 August 2010; pp. 623–629.

- 26. Kadri, A.; Yaacoub, E.; Mushtaha, M. Empirical Evaluation of Acoustical Signals for Leakage Detection in Underground Plastic Pipes. In Proceedings of the 17th IEEE Mediterranean Electrotechnical Conference, Beirut, Lebanon, 13–16 April 2014; pp. 54–58.
- 27. Ahadi, M.; Bakhtiar, M.S. Leak Detection in Water-Filled Plastic Pipes Through the Application of Tuned Wavelet Transforms to Acoustic Emission Signals. *Appl. Acoust.* **2010**, *71*, 634–639, doi:10.1016/j.apacoust.2010.02.006.
- 28. Rashid, S.; Qaisar, S.; Saeed, H.; Felemban, E. A Method for Distributed Pipeline Burst and Leakage Detection in Wireless Sensor Networks Using Transform Analysis. *Int. J. Distrib. Sens. Netw.* **2014**, *10*, 1–14, doi:10.1155/2014/939657.
- 29. Bentoumi, M.; Chikouche, D.; Mezache, A.; Bakhti, H. Wavelet DT Method for Water Leak-Detection Using a Vibration Sensor: An Experimental Analysis. *IET Signal Process.* **2017**, *11*, 396–405, doi:10.1049/iet-spr.2016.0113.
- 30. Lay-Ekuakille, A.; Vendramin, G.; Trotta, A. Robust Spectral Leak Detection of Complex Pipelines Using Filter Diagonalization Method. *IEEE Sens. J.* **2009**, *9*, 1605–1614, doi:10.1109/JSEN.2009.2027410.
- 31. Almazyad, A.S.; Seddiq, Y.M.; Alotaibi, A.M.; Al-Nasheri, A.Y.; BenSaleh, M.S.; Obeid, A.M.; Qasim, S.M. A Proposed Scalable Design and Simulation of Wireless Sensor Network-Based Long-Distance Water Pipeline Leakage Monitoring System. *Sensors* **2014**, *14*, 3557–3577, doi:10.3390/s140203557.
- 32. Shi, Y.; Zhang, C.; Li, R.; Cai, M.; Jia, G. Theory and Application of Magnetic Flux Leakage Pipeline Detection. *Sensors* **2015**, *15*, 31036–31055, doi:10.3390/s151229845.
- 33. Muggleton, J.M.; Brennan, M.J.; Pinnington, R.J. Wavenumber Prediction of Waves in Buried Pipes for Water Leak Detection. *J. Sound Vib.* **2002**, 249, 939–954, doi:10.1006/jsvi.2001.3881.
- 34. Jackets, T. The Newton-Laplace Equation and Speed of Sound. 2014. Available online: https://www.thermaxxjackets.com/newton-laplace-equation-sound-velocity/ (accessed on 20 July 2017).
- 35. Yoon, D.J.; Lee, Y.; Kim, Y.; Kim, C.; Jung, J.C.; Kim, S.M.; Lee, J.; Jeon, H.; Moon, C.; Kim, E.C. *Development of Leak Detection System for Waterworks Using Elastic Wave*; Technical Report of Korea Research Institute of Standards and Science: Daejeon, Korea, 2004; pp. 1–263.
- 36. Carter, G.C. Time Delay Estimation. Ph.D. Thesis, The University of Connecticut, Storrs, CT, USA, 1976.
- 37. Hayes, M.H. Statistical Digital Signal Processing and Modeling; John Wiley & Sons, Inc.: New York, NY, USA, 1996.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).