Multicomponent filtering for Leakage detection using bidimensional thermometric data

Amir Ali Khan, Valeriu Vrabie, Guy d'Urso, Jerome I. Mars

MULTICOMPONENT FILTERING FOR LEAKAGE DETECTION USING BIDIMENSIONAL THERMOMETRIC DATA

Amir A. Khan[†], *Valeriu Vrabie*[‡], *Guy D'Urso* [*], *and Jérôme I. Mars*[†]

[†]GIPSA-Lab–Department of Image and Signal Processing (DIS), INP of Grenoble, BP 46, 38402 Saint Martin d'Hères Cedex, France.

[‡]Centre de Recherche en STIC (CReSTIC), University of Reims, BP 1039, 51687 Reims Cedex, France.

[*]Direction R&D, Electricité de France (EDF), 6 quai Watier, 78401 Chatou Cedex, France.

Email: {amir-ali.khan, jerome.mars}@gipsa-lab.inpg.fr, valeriu.vrabie@univ-reims.fr, guy.durso@edf.fr

ABSTRACT

In the problem of leakage detection in dikes based on Distributed Temperature Sensors, different factors influence the acquired temperature. These factors having certain peculiarities in the temporal and spatial domains, the spatial and temporal gradients of the original signal can be used to construct a 3C signal. The information pertaining to the leakages, characterized by high dynamics, is present in all the three components while the information linked to other factors such as existing structures, precipitations, etc. is significant in only two components. This fact can be exploited using vectorial analysis such as vector median filtering. A vector is selected based on the criteria of minimum distance from all other vectors. A thresholding of the filtered signal residue extracts the information uniquely related to the leakages allowing their immediate detection.

1. INTRODUCTION

Multicomponent analysis has long been used for diverse applications in the domain of signal and image processing. The existence of multiple components of the same signal allows exploitation of different redundancies in the signal. The basic idea is that a certain desired information may be distributed on different components in different compositions and the ensemble processing of the various components may lead to an information which may not be so clear in each of the individual components [1, 2].

Now-a-days, an important issue in the engineering domain is the detection of anomalies, such as leakages (significant flow of water), in the dikes to avoid disaster at mass level [3]. In this regard, the use of temperature based methods presents an efficient solution, specially with the improved possibilities of temperature measurements through the use of distributed temperature sensors (DTS) based on optical fiber cables [4]. The major advantage of DTS is that they use low-cost telecommunications grade fiber while at the same time providing an ability to multiplex large number of sensors along a single fiber [5]. The problem of leakage detection and localization has not been addressed in the signal processing community before and this work addresses a completely new research problem. The idea behind the use of temperature data for leakage detection is that a change of ground temperature is brought about by significant flow of water through the structure due to leakages [6]. The leakage is characterized by the difference between air temperature and water temperature in canal. Thus, in case of leakage, the temperature acquired by DTS is directly proportional to this temperature difference. It also depends on the leakage flow rate as well as the distance of the actual leakage from DTS. The farther the leakage from the sensors, more pronounced will be the effect of external factors such as wind speed, etc. However, this change of temperature can equally be brought about by other factors such as the seasonal variations, precipitation, the existing structures (e.g. drains), etc.

Each of the above mentioned factors have certain peculiarities. While some of them are localized in time (precipitation, seasonal variations), others are localized in distance (drains, etc.) while the interesting factors such as the leakages are localized both in time and in distance. In this paper, we utilize these temporal and spatial redundancies of the acquired temperature signal for the purpose of leakage detection. We present that how using this mono-component acquired temperature signal, we can create a three-component (3-C) temperature signal. Then, we show how, using the techniques of vector median filtering, usually employed in color image processing [7, 8], on this 3-C signal, we can exactly detect the presence of leakages in the dikes. The obtained results for the 3-C case are compared with both one-dimensional and bidimensional median filtering for a real temperature data set.

2. TEMPERATURE DATA AND ITS MULTICOMPONENT EXTENSION

The temperature data is acquired over the 2km length of the optical fiber, laid close to a canal of a hydroelectric center, with a sampling rate of 1m and to keep track of the temporal evolution of various phenomena, numerous signals (or profiles) are acquired over time with a 2hr interval during the months of April and May. This gives a two-dimensional temperature signal, $y(x,t)$, as a function of displacement along the fiber and time and the recorded data set can be written in a matrix format as:

$$\mathbf{Y} = \{y(x,t) \mid 1 \leq x \leq N_x, 1 \leq t \leq N_t\}, \tag{1}$$

where N_x and N_t are the number of observation points and the total time of acquisition, respectively.

Three artificial leakages, of percolation type, were introduced during different times in the month of May with different flow rates and positions over fiber. Also a hot point (an artificial flow of hot water) was tested. A description of these leakages along with the localization of some existing drains (cement structures) present in the path of the optical fiber is given in Tab. 1 whereas the data itself is given in Fig.1(a).

Table 1. Characterization of the structures and leakages for the site Oraison

	Leakages			Hot Point	Drains	
	L1	L2	L3	HP	D1	D2
Location (km)	1.562	1.547	1.569	0.674	0.561	0.858
Time (day)	39(noon)	41(morn)	41(eve)	39(morn)	-	-
Flow rate (lit/min)	5	1	1	-	-	-

The temperature signal incorporates also other phenomena such as precipitations, response of the near surface where the acquisitions are made, seasonal variations, etc. The inherent characteristics of all these factors are their localization in distance/time. If we consider the case of precipitations, they are in general well localized in time but slowly varying in distance. Likewise, the structures such as drains are localized in distance and exist for all times. Meanwhile, the factors of our interest, the leakages, are localized both in time and in distance. Thus using gradient operator will exploit these inherent characteristics.

For the temporal gradient, the gradient operator is applied for each sensor along the entire time duration of the signal whereas in the spatial gradient, the gradient operator is applied for each profile along the entire length of the fiber. The former thus attenuates the effects of existing structures like drains (which exist for all times) and seasonal variations, while the latter attenuates the effects of the ground response as well as that of the precipitation (which vary slowly along the distance). We can thus construct a 3-C signal in the form of a data cube:

$$\mathcal{Y} = \left\{\mathcal{Y}(:,:,1) = \mathbf{Y}, \mathcal{Y}(:,:,2) = \frac{\partial \mathbf{Y}}{\partial x}, \mathcal{Y}(:,:,3) = \frac{\partial \mathbf{Y}}{\partial t}\right\} \tag{2}$$

This multicomponent data is represented in Fig. 1. The vertical line in Fig. 1(c) represents the interpolation done for the three day non-acquired data.

3. MULTICOMPONENT FILTERING

If we consider the data cube constructed previously, we would notice that at each distance, x, and time, t, we have a vector, $\mathbf{y}_i = \mathcal{Y}(x,t,:) \in \mathbf{R}^3$. These vectors at different times and distances would at most contain two significant elements. This would be true for all times and distances except at the instants and positions of the leakages. Here, we would have all the three significant elements. Thus, the distance between this vector and others in its vicinity would be very large compared to distances computed in other zones. We can thus devise a vector based filtering measure which looks to exploit this very fact. One of the measures can be the vector median filtering (VMF) operator which was initially proposed as an extension of the scalar median filtering for impulsive noise removal [8]. The VMF can be derived with vector order-statistics techniques using the Minkowski metric, to quantify the distance between two vectors in the magnitude domain [7]. The methodology can be formulated as in the following equation:

$$\mathbf{y}_{min} = \arg\min_{\mathbf{y}_i \in \mathbf{W}} \sum_{j=1}^{N} \|\mathbf{y}_i - \mathbf{y}_j\|_2, \tag{3}$$

where $\mathbf{W}$ is a filtering window containing the N vectors to be compared, $\mathbf{y}_i$ (or $\mathbf{y}_j$), $\mathbf{y}_{min}$ is the vector selected amongst the N vectors in the processing window and $\|.\|_2$ is the L_2 vector norm known as the Euclidean distance.

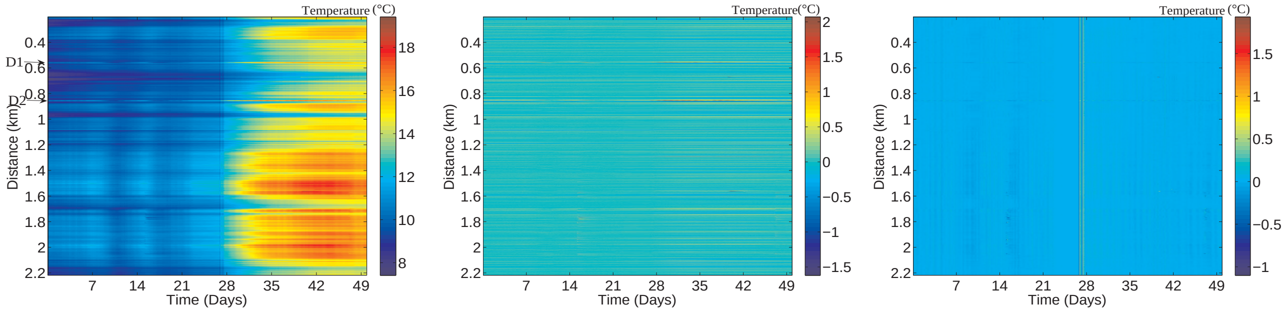
The acquired temperature signal contains three leakages and a test of hot point, all of which have approximately impulsive signatures (information) in both time and distance as opposed to the other processes in the signal. The effect of applying a VMF scheme would therefore be to annihilate this information. As a result, the residue obtained as a difference between the initial data set and the filtered data set should contain the information mostly linked to these processes.

4. APPLICATION ON TEMPERATURE DATA

The VMF described in the last section is applied on the 3-C signal, described in Sec. 2 with a processing window, $\mathbf{W}$, of size 3×3, giving $N = 9$ vectors. This window size was chosen because it translates to 3m long window with a 6hr duration (in terms of current acquisition parameters), thus incorporating smaller dynamics both in time and distance except for instances of leakages and possibly precipitations. The resulting filtered data, $\mathcal{Y}_{filtered}$, is obtained by sliding the processing window to scan all positions (x) and instants (t) and replacing the center of each window by the result of Eq. (3). We can thus obtain the residue on the first component as:

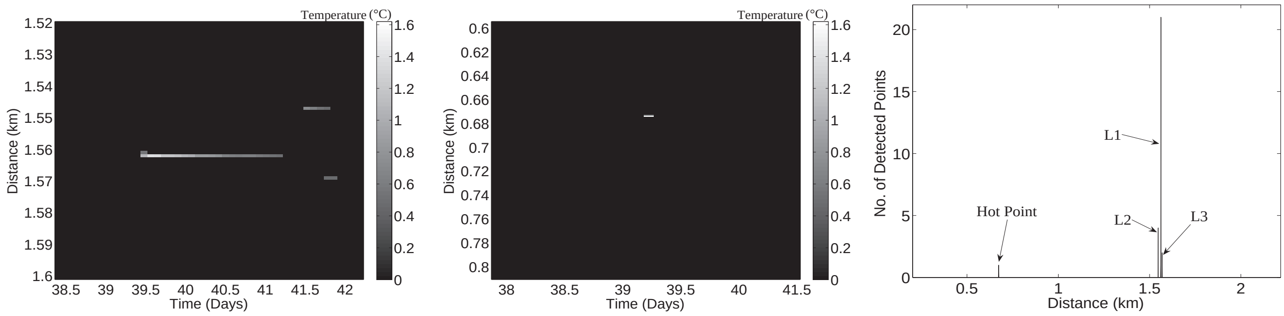
$$\mathbf{Y}_{residue} = \mathbf{Y} - \mathbf{Y}_{filtered}, \tag{4}$$

where $\mathbf{Y}$ is the acquired data and $\mathbf{Y}_{filtered}$ is the first component of $\mathcal{Y}_{filtered}$ (i.e. $\mathcal{Y}_{filtered}(:,:,1)$).



(a) Original Temporal Data (b) Spatial Gradient (c) Temporal Gradient

Fig. 1. Multicomponent Temperature Data from Monocomponent Data.



(a) Zoom in vicinity of Leakages (b) Zoom in Hot point vicinity (c) Projection $P(x)$

Fig. 2. Detection of Leakages in VMF Residue.

The residue thus obtained contains pre-dominantly the information linked to the leakages. To better detect the presence of leakages in the residue, a thresholding is applied. The value of this threshold was determined based on the physical aspect of temperature detection by DTS. The leakage temperature is directly proportional to the difference of temperature between that of air and water of the canal as well as to the leakage flow rate. The effect of external factors such as wind speed can be neglected in the present case as the percolated leakages are closer to the actual sensors. The current site is controlled in terms of leakage positions and flow rates and it was observed that a low flow rate leakage (1 lit/min) in a time zone where the difference of temperature between that of air and water is small (evening) brings a change of $0.51°C$ and thus the threshold is set at $0.5°C$. The resulting thresholded residue is:

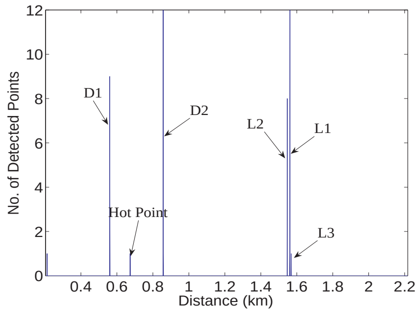
$$y^{th}_{residue}(x,t) = \begin{cases} y_{residue}(x,t) & \text{if } y_{residue}(x,t) > \text{threshold} \\ 0 & \text{otherwise} \end{cases}$$

A zoom of the thresholded residue in the zone around leakages and the hot point shows that the residue is perfect for the purposes of leakage detection (Fig. 2(a)-(b)). The same can be observed by seeing the projection,

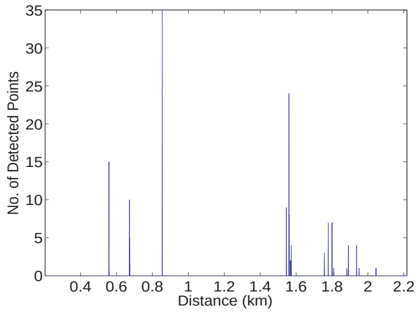
$$P(x) = \sum_{t=1}^{N_t} y^{th}_{residue}(x,t), \tag{5}$$

of the thresholded residue in Fig. 2(c). This projection reveals clearly that the three leakages and the hot point are the only processes detected in the residue. The leakages are detected at exactly the same distances as in the Tab. 1, whereas the residue is devoid of all other phenomena which were present in the initial acquired temperature signal thus giving a perfect identification.

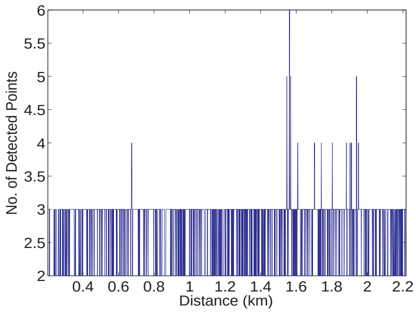
In order to highlight the advantages of the multicomponent analysis as compared to the monocomponent one, we tested the scalar median filtering. The 2-D median filtering results for the 3×3 window size using the same thresholding as above reveals neither the hot point nor the leakage $L3$ and the detection is thus incomplete. Resetting the threshold manually in order to keep the leakage $L3$ in the residue, the results in Fig. 3 are obtained. It can be observed that the residue incorporates undesired detection of drains (Fig. 3(a)), their detection being linked to the fact that they are affected by the day/night temperature variations being more exposed to the air. The application of 2-D median filtering on spatial gradient (Fig. 3(b)) results in undesired detection of drains, less accurate localization of the hot point and the leakages as well as putting into effect some false phenomena at positions around 1.8km. The results of temporal gradient are completely inexploitable (Fig. 3(c)) in the sense that setting the threshold with respect to the weakest detected leakage, $L3$, means that separation of leakages from the rest of the data is impossible.



(a) Projection $P(x)$ for $\mathbf{Y}$



(b) Projection $P(x)$ for $\frac{\partial \mathbf{Y}}{\partial x}$



(c) Projection $P(x)$ for $\frac{\partial \mathbf{Y}}{\partial t}$

Fig. 3. Results of 2D Median Filtering on individual components for 3×3 window.

The 1-D median filtering applied for each sensor (row-wise) does not give any considerable results. The results obtained by applying the filtering for each acquisition (column wise) of the original data show that the detected leakages are corrupted (Fig. 4) by the drains that are quite dominant in the residue and the hot point is not even detected. Moreover, there are certain unknown points identified in the residue, some of which can be attributed to the spatial zones more affected by precipitation. The application of 1-D filtering on the other two data components give highly unsatisfactory results. Other dimensions of the processing window were also tested for both 2-D and 1-D filtering not giving better results either. This shows the efficiency of multicomponent filtering in terms of its capability to detect only the leakages and the hot point while in addition allowing to set a threshold based on the physical acquisition principle.

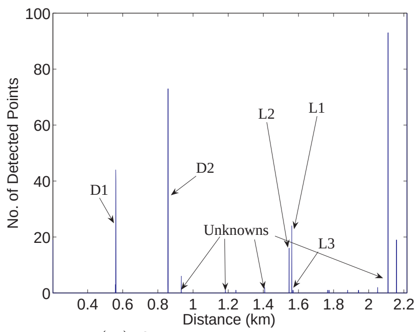


Fig. 4. Projection $P(x)$ for column-wise 1D Median Filtering on $\mathbf{Y}$ for window size 3.

5. CONCLUSION

In this paper, a method for the percolation type leakage detection in dikes using DTS data was proposed and tested on a real temperature data set. The acquired temperature data contains information linked to leakages but this information is largely mixed with other phenomena such as the drains, seasonal variations, precipitations, response of the near surface, etc. The spatial and temporal gradients help remove the influence of all the phenomena which vary slowly with distance (like the response of the near surface and precipitations) and time (like the drains and the seasonal variations), respectively. A 3-C signal can thus be constructed using the original temperature signal and its temporal and spatial gradients. The leakages being well localized both in time and in distance, are removed by applying a vector median filtering to this 3-C dataset. The resultant thresholded residue contains the information solely linked to them thus providing a very efficient means of leakage detection. In this way, we succeed in showing how the information not immediately exploitable in the individual monocomponents can be put to evidence using the vectorial filtering techniques on the multicomponent signal.

6. REFERENCES

[1] C. Paulus, J. I. Mars, and P. Gounon, “Wideband spectral matrix filtering for multicomponent sensors array,” *Signal Processing*, vol. 85, pp. 1723 – 1743, Sept. 2005.

[2] J. I. Mars, N. L. Bihan, V. Vrabie, and C. Paulus, “Multicomponent seismic data filtering by multilinear methods,” in *Proc. ISCCSP*, Morocco, 2006.

[3] S. Johansson, “Localization and quantification of water leakage in ageing embankment dams by regular temperature measurements,” in *Proc. ICOLD, 17th Congress*, Vienna, 1991.

[4] S. Yin, “Distributed fiber optic sensors,” in *Fiber Optic Sensors*, F. T. Yu and S. Yin, Eds., pp. 201–229. Marcel Dekker Inc., 2000.

[5] A. H. Hartog, “A distributed temperature sensor based on liquid-core optical fibers,” *IEEE Journal of Lightwave Tech.*, vol. 1, pp. 498–509, Sept. 1983.

[6] B. Vogel, C. Cassens, A. Graupner, and A. Trostel, “Leakage detection systems by using distributed fiber optical temperature measurements,” in *Proc. SPIE Smart Structures and Materials*, D. Inaudi E. Udd, Ed., 2001, vol. 4328, pp. 23–34.

[7] R. Lukac, B. Smolka, K. Martin, K.N. Plataniotis, and A.N. Venetsanopoulos, “Vector filtering for color imaging,” *IEEE Signal Processing Magazine*, vol. 22, pp. 74 – 86, Jan. 2005.

[8] J. Astola, P. Haavisto, and Y. Neuvo, “Vector median filters,” *Proc. IEEE*, vol. 78, pp. 678 – 689, April 1990.