ISSN 1726-5479

SENSORS & TRANSDUCERS

vol. 238
11/19

Sensors and Transducers: Design & Applications II

Electronic version. Not for distribution.

20 years IFSA

International Frequency Sensor Association Publishing

Sensors & Transducers

International Official Open Access Journal of the International Frequency Sensor Association (IFSA) Devoted to Research and Development of Sensors and Transducers

Volume 238, Issue 11 November 2019

IFSA Publishing: Barcelona • Toronto

Published by International Frequency Sensor Association (IFSA) Publishing. Printed in the USA.

Sensors & Transducers

Volume 238, Issue 11
November 2019

www.sensorsportal.com

e-ISSN 1726-5479
ISSN 2306-8515

Sensors & Transducers journal is an open access, peer review international journal published monthly by International Frequency Sensor Association (IFSA).
Available in both: print and electronic (printable pdf) formats.

Sensors & Transducers Journal

Contents

Volume 238
Issue 11
November 2019

www.sensorsportal.com

ISSN 2306-8515
e-ISSN 1726-5479

Research Articles

Authors are encouraged to submit article in MS Word (doc) and Acrobat (pdf) formats by e-mail: editor@sensorsportal.com. Please visit journal's webpage with preparation instructions: https://www.sensorsportal.com/HTML/DIGEST/Guides_for_Authors.htm

International Frequency Sensor Association (IFSA).

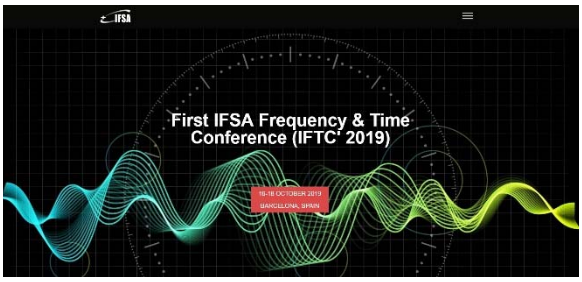

Sensors Industry News
ALL about SENSORS
FREE Monthly IFSA Newsletter
ISSN 1726-6017
SUBSCRIBE NOW
subscribe@sensorsportal.com


SENSORS WEB PORTAL
IFSA
MEMS
NEMS
NANOSENSORS
SMART SENSORS
Sensors Web Portal
All about SENSORS
http://www.sensorsportal.com

Foreword to Special Issue

Connecting Academy and Industry
(IFSA slogan)

It is my great pleasure to introduce the 2nd volume of special issue of *Sensors & Transducers* journal, Vol. 238, Issue 11, November 2019, entitled *'Sensors and Transducers: Design & Applications II'*. This special issue contains 14 selected and extended papers from the *1st IFSA Frequency & Time Conference (IFTC' 2019)* and *5th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2019)*. This conference umbrella has taken place on 25–27 September 2019 in Adeje, Tenerife (Canary Islands) Spain.



IFTC' 2019 Conference Web Site.

Photo 1. Plenary session.

Photo 2. Poster session.

Photo 3. Closing session.

The events were organized by *International Frequency Sensor Association (IFSA)* – one of the major professional, non-profit association serving for sensor industry and academy more than 20 years.

The *IFSA* has been founded on 23 August 1999 with the aim to encourage the research and development in the area of sensors with frequency, time interval, duty-cycle and digital output, including multifunction, multifrequency transducers, the usage of new physical effects and principles. *IFSA* reports on current research in this field worldwide through its media: Sensors Web Portal (http://www.sensorsportal.com), IFSA Newsletter, Sensors & Transducers journal, e-Digest (magazine), monographs, open access books and related conferences. Today *IFSA* has more than 750 members from 71 countries from academia and industry worldwide. Among them are: *ABB, ams, Analog Devices, Bosch, Cadence, Caterpillar, Freescale, General Electric, Honeywell, Huawei, imec, John Deere, Mazda, Meggitt, Mentor, Philips, Tata, Texas Instruments, TDK, Yokogava* and other industrial (60 %) and academic (40 %) members.

Twenty years *IFSA* activities has been reported at the conference's panel discussion '*Sensor Community: which services do we need ?*' with the following panelists and keynote presenters: Boris Farber (*Zircoa, Inc.*, USA), Simona M. Cristescu (*Radboud University*, The Netherlands), Carlos Ruiz Zamarreño (*Public University of Navarra*, Spain), José Miguel Dias Pereira (*Polytechnic Institute of Setúbal*, Portugal) and Dan Mihai Stefanescu (*Romanian Measurement Society*, Romania), moderated by Sergey Y. Yurish (*IFSA*, Spain). Prof. José Miguel Dias Pereira has also presented the IFTC' 2019 conference keynote presentation entitled '*A/D*

Conversion Techniques Based on the Usage of Pulse Width Modulated (PWM) Signals: Applications for Digital Sensors and Sensor Systems'.

I would like to thank all authors for submitting their latest work, thus contributing to the excellent technical contents of the journal.

A limited number of the best articles from this special journal issue will be selected for extension into book chapters for publication in our popular open access Book Series '*Advances in Sensors: Reviews*', Vol. 8 or '*Advances in Measurements and Instrumentation: Reviews*', Vol. 2, which will be published in 2020/2021. Through the journal and book open access, the reach and impact of obtained and reported at the IFTC' 2019 and SEIA' 2019 conferences research results will be substantially increased.

The *2nd IFSA Conference on Frequency & Time (IFTC' 2020)* will take place on 23-25 September 2020 in Porto, Portugal in conjunction with *the 6th International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2020)* under the same conference umbrella.

I hope you will enjoy this journal issue !

Sergey Y. Yurish
Editor-in-Chief

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 1-7

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Data Analysis of Frequency-analog Signals of a Humidity Sensor to Improve the Detection of Fakes

Anett BAILLEU
HTW Berlin – University of Applied Science, FB I – Energy and Information,
Wilhelminenhofstr. 75 A, 12459 Berlin, Germany
Tel.: +49 30 5019 3341
E-mail: anett.bailleu@htw-berlin.de

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: The aim of the described solution is to detect the vitality of presented measurement objects. Therefore, a capacitive sensor is used. The output signal of the sensor is a frequency. The frequency-analog output signal is depending on the humidity of the measurement object. The sampled frequencies data, monitored like dynamic time series data, have been analyzed to develop an efficient binary classifier. The classifier functions as a vitality checker. Vitality checkers are useful additional tools in the field of biometric applications to improve the security of such systems.

Keywords: Analysis of frequency-analog time series data, Vitality checker, Binary classifier.

Preface

This is an extended paper of a work which was discussed at the first IFSA Frequency & Time Conference (IFTC´2019) [1]. It is supplemented by some specific results of analysing the scattering and the temporal change of the time series data.

Whereas in [1] representative qualitative time series of period values were published, in this paper raw data plots in form of quantitative frequency time series are shown.

1. Introduction

Biometric systems have been gaining in importance for more than two decades. They have been used in different areas of social and economic life, in particular for access control to secured systems. The detection of biometric features works very well in the present. So, biometric systems, especially fingerprint sensors, are very often used for identification and authentication purposes. Especially biometric systems, that meet the challenges of sovereign, protected security areas and the growing security demands of privately run businesses, require constant improvements. The detection of fakes is one of the main focal points of such improvements. An almost absolute reliable control mechanism to differ a real biometric object from a fake is to check its characteristic to be alive. On the one hand, it is possible to produce identical copies of biometric features, but on the other hand, it is nearly impossible to imitate the vitality of a measurement object in a sufficiently realistic way. That is why, vitality checkers are components of modern sophisticated biometric systems [2-4], e.g. in sovereign, protected security systems [5].

In this context, a lot of ideas to check the vitality were developed and published [6-14]. Aside from the measurement of vitality parameters like heart rate and oxygen saturation, the verification of the peripheral blood flow, the detection of temperature and patterns of movements etc. have been discussed so far.

Different from these previous works, in this paper a solution based on data analyses of the signals of a frequency-analogue humidity sensor is described.

2. Solution Approach

Water is a permanent metabolic product of human beings. This effects a measurable ambient humidity of skin, e.g. due to the activity of the sweat glands. State of the art vitality checkers using the activity of the sweat glants are based on image processing solutions [15-16].

Here, a solution approach is shown, that is based on the idea of data analyses of time series signals of a capacitive humidity sensor. The studies do not target on very accurate measurements of humidity.

In a multitude of test cases, a binary classifier works efficient to separate real measurement objects from fakes by a measurement signal which correlates with the humidity.

It suggests that dry measurement objects are most probably fakes and wet measurement objects are very often real biometric objects, which are alive. Furthermore, it hypothesized, that adequately moistened fakes will be classified correctly in most cases by using appropriate methods of data analyses. Therefore, several data analyses methods to classify time series signals are used.

3. Experiments

3.1. Experimental Setup

The most important hardware component of the measurement chain is the primary sensor, because it determines the functional characteristic of the whole solution. The primary sensor is a capacitive humidity sensor. The capacity between interdigital electrodes depends on the ambient humidity in the electrical stray field of the used capacitor. Fig. 1 shows a greatly enlarged photo of the used sensor. It is a dimensioned picture.

The dimensions of the sensor are round about 3 mm × 3 mm. Because of its small dimensions, it is not problematic to integrate this sensor in biometric systems.

In the center is the relevant sensitive area of 1.7 mm × 1.7 mm. There, the planar interdigital electrodes are located.

Fig. 1 includes a zoom of the layout of the planar electrodes. The strips of the electrodes and the gap between the electrodes amount to only a few micrometers (≤ 5 µm).

The output signal of the sensor is converted into a square-wave signal by an integrated electronic based on the 555-timer-circuit (Fig. 2), internally consisting of two comparators and a RS-flipflop.

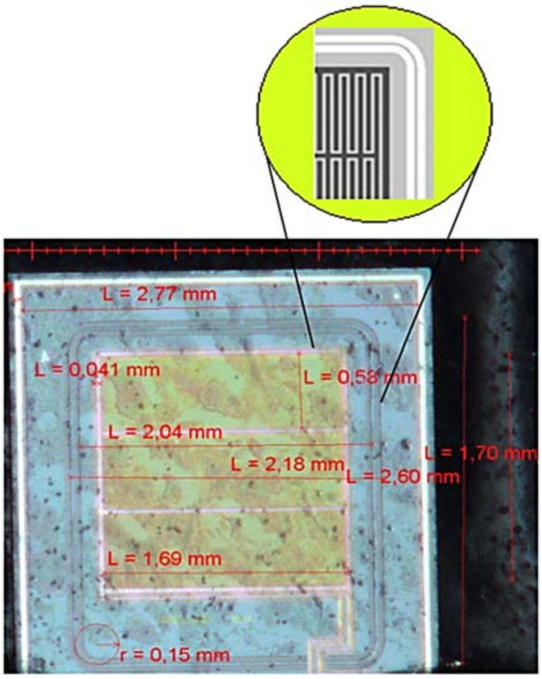


Fig. 1. Greatly enlarged photo of the capacitive humidity sensor (with schematic zoom of planar interdigital electrodes).

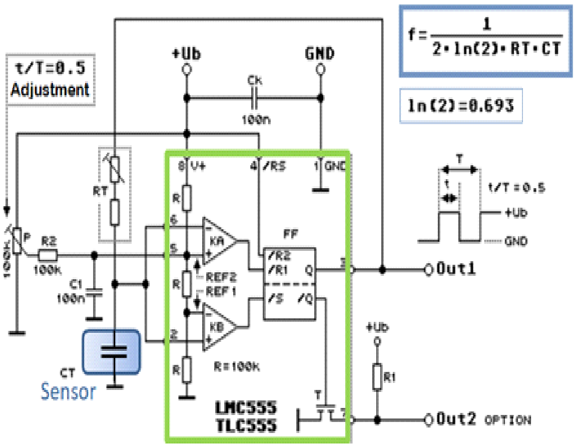


Fig. 2. Capacitance to frequency converter.

The frequency depends directly on the change of capacity caused by the effects of ambient humidity. The square-wave signal has been sampled with a suitable data acquisition device (DAQ), specified by an ADC resolution of 16 Bit and a sample rate of 200 kS/s.

The DAQ has been connected via USB with a computer.

Fig. 3 illustrates the complete measurement chain.

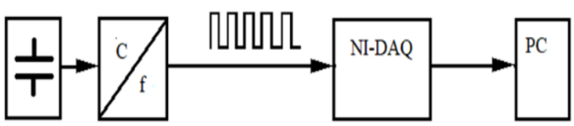


Fig. 3. Measurement chain.

The average frequency respectively the average period of the square-wave signal can be calculated by counting of zero passages. This calculation (Eq. (1)) has been done in timeslots of 7.5 ms:

$$f_i = \frac{N}{7{,}5\,\text{ms}},\; T_i = \frac{1}{f_i} = \frac{7{,}5\,\text{ms}}{\text{N}}, \qquad (1)$$

where f-frequency, T-period, i-count-index: 0…799.

Then, the average values are stored in a raw-data-set.

3.2. Test Objects

The test objects were real alive fingers of more than 70 persons and 180 fakes. The 180 fakes consist of 124 representative dry fake-examples and 56 representative wet fake-examples. The used test-fakes represent the known state of the art of fakes, but details to produce them will not be described in this paper.

While dataset 1 contains only the measurement results of index fingers of the right hand of different persons, dataset 2 contains results of measurements of all ten fingers of both hands from the test persons. In dataset 2 included are repeated measurements of the same fingers at several times. There are more raw-data-sets of index fingers than of the other eight fingers in the dataset 2. It should be noted, that in the presented diagrams, the results of measurements with real alive fingers are called 'Real', all the other ones are called 'Fake'.

4. Results

First, in this section selected raw-data-plots are shown, because significant differences can be observed in these diagrams. Then, in a separate section, results of several data analysis procedures are illustrated.

4.1. Raw-Data-Plots

Figs. 4-6 present representative time series of calculated frequency values (see Eq. (1)). They show the change of the frequency-values in a limited duration.

The frequency values between 55 kHz and 57 kHz of a real alive finger (Fig. 4) are significantly lower than those of dry fakes between 133 kHz and 135 kHz (Fig. 5).

But the frequencies of real fingers and wet fakes may be nearly the same (compare Fig. 4 and Fig. 6).

Furthermore, the diagram quality of real vital objects seems to be 'noisy with a drift', while the diagram quality of the fakes (dry and moist) seems to be 'square-quantized'.

The visual distinguishability of the diagrams is a strong indicator to find a technical opportunity for classifying these objects (Figs. 4-6).

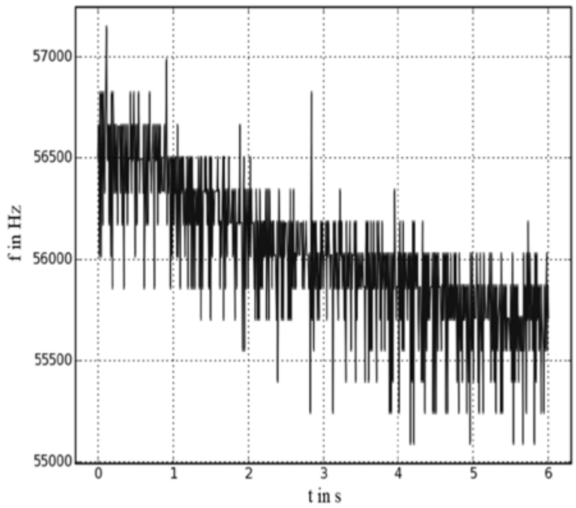


Fig. 4. Frequency-time series of the humidity sensor signal generated by measuring at a real alive index finger.

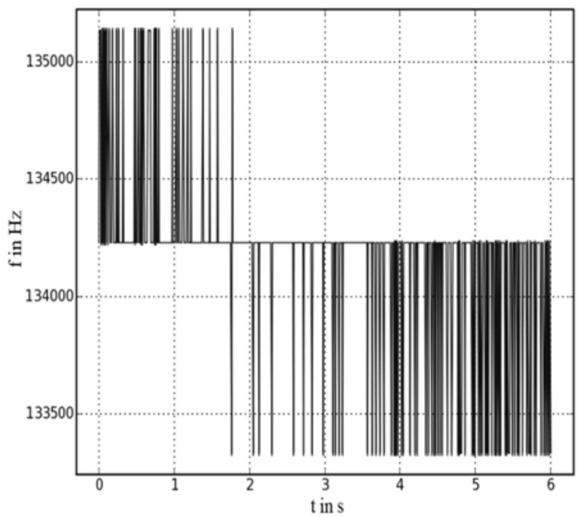


Fig. 5. Frequency-time series of the humidity sensor signal generated by measuring at a dry fake.

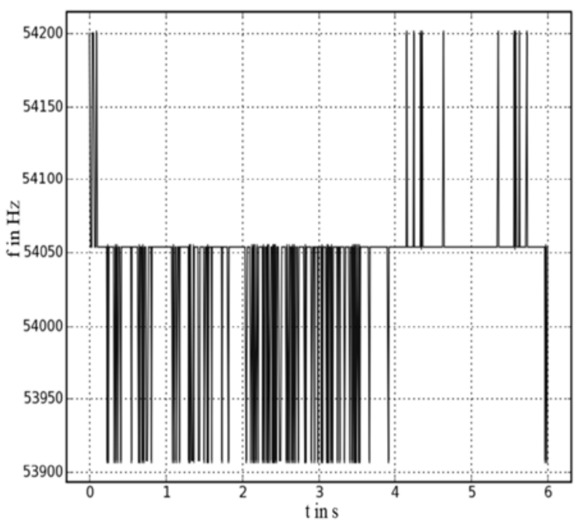


Fig. 6. Frequency-time series of the humidity sensor signal generated by measuring at a wet fake.

4.2. Results of Classification

4.2.1. Frequency-threshold as Separation Criterion

The measured signals of different objects had to be evaluated to develop a reliable binary classifier to separate the real alive biometric objects from fakes.

Typical quantitative criterions to characterize the quality of a classifier are:

$$specifity = \frac{TN}{TN+FP} = \frac{TN}{N}, \quad (2)$$

$$sensitivity = \frac{TP}{TP+FN} = \frac{TP}{P}, \quad (3)$$

$$accuracy = \frac{TP+TN}{TP+FN+TN+FP} = \frac{TP+TN}{P+N}, \quad (4)$$

$$precision = \frac{TP}{TN+FP} = \frac{TP}{N}, \quad (5)$$

$$Youdenindex = \frac{TP}{P} + \frac{TN}{N} - 1, \quad (6)$$

where P is the positive (real alive finger), N is the negative (fakes), TN is the true negative, TP is the true positive, FN is the false negative, FP is the false positive.

That the frequency respectively the period of the square-wave signal of the humidity sensor are expedient separating criterions to detect dry fakes has been illustrated in Fig. 7.

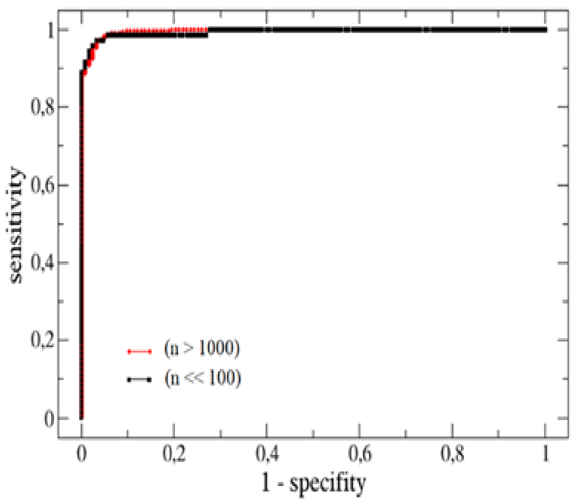


Fig 7. Frequency value as separation criterion (ROC-curves for real fingers vs. dry fakes).

Receiver-Operating-Characteristics (ROC-curves) are shown for detection of dry fakes vs. the real fingers of dataset 1 and dataset 2. Both ROC-curves are nearly identic.

The capacitive humidity sensor is sensitive enough to detect the thin sweat film on the skin of the real fingers.

Ca. 97 % correct classification results for both classes are obtained if only dry fakes are used.

Fig. 8 shows a combination of diagrams to find out the optimal frequency-threshold for separation of real fingers from all kind of suitable fakes.

ROC-curves for real fingers vs. fakes on the left side and Youdenindex-curves and histograms on the right side are shown in Fig. 8. On top are results of dataset 1 vs. all fakes, further below are results of dataset 2 vs. all fakes.

By using wet fakes the performance decreases (Fig. 8), but is good enough by using only one frequency threshold as separating criterion.

Overlaps in both histograms are found for all frequency thresholds in Fig. 8. The best classification results are found for a frequency-threshold at the maximum of the Youdenindex at 0.66 in this study (Eq. (6)). A sensitivity of 96.7% (Eq. (3)) and a specifity of 69.4 % (Eq. (2)) are reached for a threshold of 134 kHz.

To improve the performance, it has to be searched for other separating criterions.

4.2.2. Separation by Scattering of Time Series Signals

It was observed, that the variance of time series values of vitality objects is higher than the variance of dry and wet fakes. Measurement results of real alive fingers were described as ‘noisy with a drift’, the non-vitality objects were characterized as ‘square-quantized’.

The dot plots in Fig. 9(a), Fig. 9(c) illustrate the difference of scattering of both classes.

So, the usability of standard deviation of the time series values as a separation criterion was tested. The results of this idea are plotted in the ROC-curves in Fig. 9(b), Fig. 9(d).

According to expectation, the ROC-curves of both datasets are very similar, like in Fig. 9. But, it seems to be worsening compared with the ROC-curves in Fig. 9.

An expedient separation threshold for the standard deviation couldn’t be found.

4.2.3. Separation by Temporal Change of Time Series Signals

Another difference between vital fingers and non-vital fakes are the temporal change in the raw data plots. The derivation of standard deviation was calculated to characterize the temporal change and a ROC-curve with this parameter as a separation criterion is shown in Fig. 10.

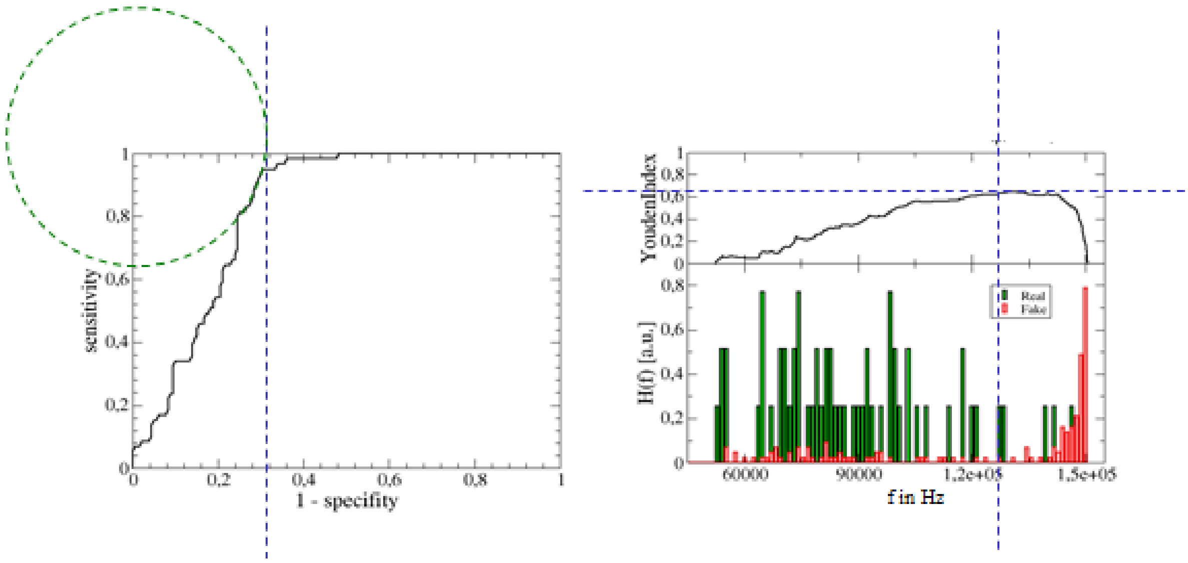


Fig. 8. Frequency threshold in ROC-, YOUDENindex-curves and histograms.

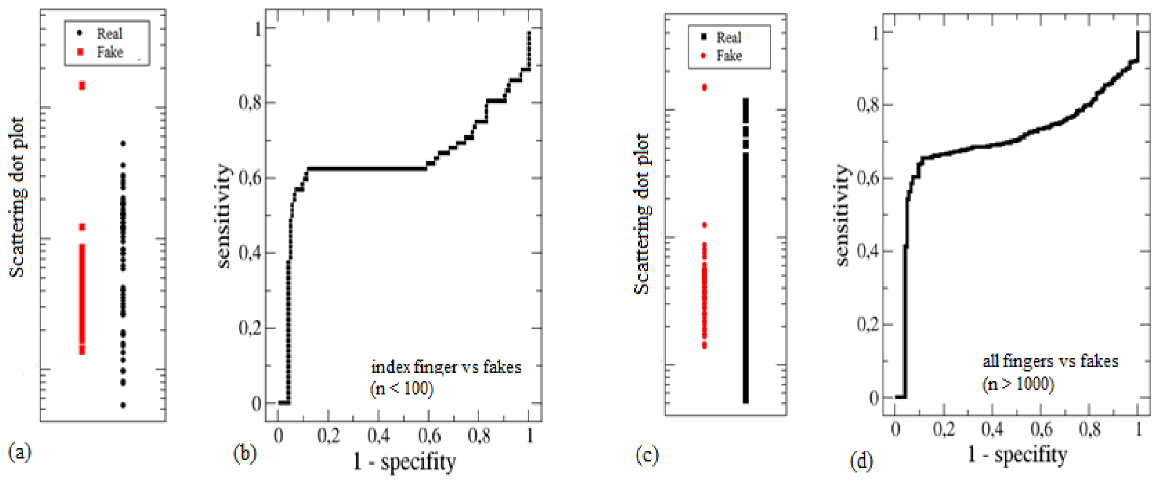


Fig. 9. Standard deviation as separation criterion (Scatter dot plots and ROC-curves.

Approximately 80 % correct classification results for real fingers and all suitable fakes seems to be attainable by using an optimal threshold of the derivation of standard deviation.

The plot in Fig. 11 was created to combine the differences of scattering and drift in one diagram. The result of using the standard deviation and the derivation of standard deviation for dataset 1 are plotted in a full-logarithmic diagram. This idea offers an interesting opportunity for a binary classification.

The line, which is printed into the diagram, shows one successful separation opportunity by combining scattering and drift characteristic.

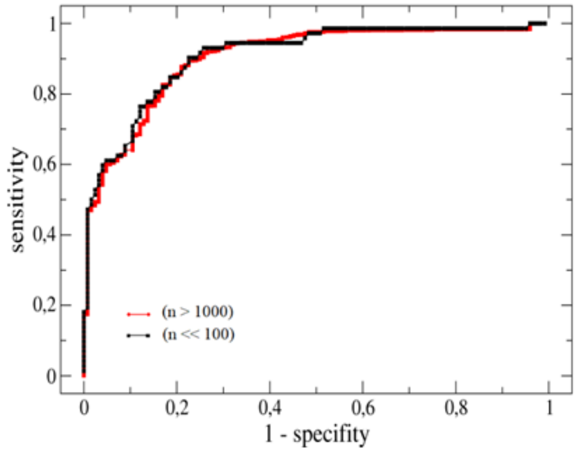


Fig. 10. Derivation of standard deviation as separation criterion (ROC-curves).

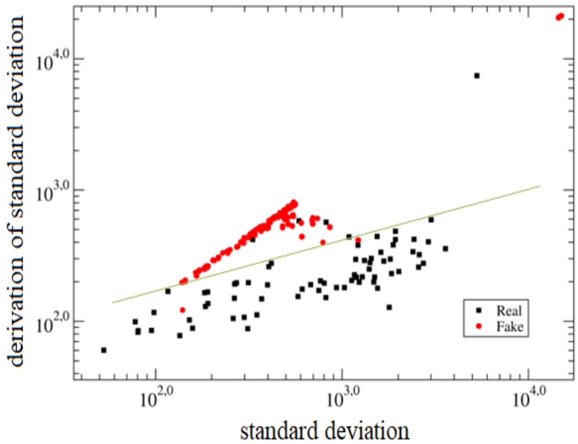


Fig. 11. Separation opportunity by using standard deviation and derivation of standard deviation (log-log-plot).

5. Conclusion

The vitality of presented measurement objects can be detected in a relative simple way by an additional capacitive sensor and the analysing of frequency-output signals. An efficient separating criterion for a binary classifier is a frequency threshold (Figs. 7, 8), but also other separating criterions based on the measurement of the output frequency value and its deviation can be found (e.g. Figs. 10, 11). That's why the described solution is suggested as an additional tool in biometric systems, especially in fingerprint sensors.

By fusion of more than one solution a classification success of more than 95 % seems to be a realistic perspective.

For the reasons mentioned, an implementation of such a relatively simple hardware solution with a planar humidity sensor can be highly recommended for a sensitive detection of fakes in biometric systems like fingerprint sensors. Furthermore, the implementation of other simple coexisting sensors is concurrently recommended to improve the detection of fakes in biometric systems. To create such systems for generic classification tasks by using commercial sensors is one of the research and development fields in the measurement laboratory at the HTW Berlin [17].

The hardware fusion of multifunctional sensors on a chip is an actual trend for more than ten years. So, it may be predicted that such chips [18] are more and more suitable on the market of electronic components.

Acknowledgment

I give big thanks to Thomas Henning for his important input in a lot of relevant scientific discussions. Mr. Henning played a determining part for the success of our project, he realized a lot of data analysing software tools.

References

[1]. A. Bailleu, Frequency-analogue Working Vitality Detector, in *Proceedings of the First IFSA Frequency & Time Conference*, 2019, pp. 9-12. ISBN 978-84-09-14539-3.

[2]. T. Matsumoto, H. Matsumoto, K. Yamada, S. Hoshino, Impact of Artificial Gummy Fingers on Fingerprint Systems, in *Proceedings of the SPIE*, Vol. 4677, 2002, pp. 275-289.

[3]. I. Goicoechea-Telleria, R. Sanchez-Reillo, J. Liu-Jimenez, R. Blanco-Gonzalo, Attack Potential Evaluation in Desktop and Smartphone Fingerprint Sensors: Can They Be Attacked by Anyone?, *Wireless Communications and Mobile Computing*, Vol. 2018, Article ID 5609195, 2018, pp. 1-16. https://doi.org/10.1155/2018/5609195.

[4]. B. Fernandez-Saavedra, R. Sanchez-Reillo, R. Alonso-Moreno, C. Sanchez-Avila, Evaluation methodology for fake samples detection in biometrics, in *Proceedings of the 42nd Annual IEEE International Carnahan Conference on Security Technology*, 2008, pp. 233–240.

[5]. https://de.newskitchen.eu/2018/08/23/%E2%80%A9bsi-erteilt-common-criteria-zertifikat-fur-dermalog-fingerabdruckscanner/ [23.08.2018]

[6]. G. L. Marcialis, *et al*., Large Scale Experiments on Fingerprint Liveness Detection, in *Proceedings of the Joint IAPR Int. Work. on Structural and Statistical Pattern Recognition (SPR & SSPR 2012)*, Hiroshima, Japan, 7-9 November 2012, *Springer*, LNCS 7625, 2012, pp. 501-509.

[7]. Biometrics: Enhancing Security or Invading Privacy?, *The Irish Council of Bioethecs*, Dublin, 2009.

[8]. A. A. Hassan, M. B. Ahmad, Enhancing the Security of Biometric Systems on View of BioFM, in *Proceedings of the International Conference on Communications and Information Technology (ICCIT)*, 2012, pp. 77-100.

[9]. G. Arunalatha, M. Ezhilarasan, Fingerprint Spoof Detection Using Quality Features, *International Journal of Security and its Applications*, Vol. 9, Issue 10, 2015, pp. 83-94.

[10]. M. Sanström, Liveness Detection in Fingerprint Recognition Systems [Diploma thesis], *Linköping University*, Linköping, Sweden, 2004.

[11]. M. Drahansky, M. Dolezel, J. Vana, E. Brezinova, J. Yim, K. Shim, New Optical Methods for Liveness

Detection on Fingers, BioMed Research International, Vol. 2013, Article ID 197925, 2013, pp. 1-11. https://doi.org/10.1155/2013/197925.

[12]. M. Drahanský, W. Funk, R. Nötzel, Liveness detection based on fine movements of the fingertip surface, in *Proceedings of the IEEE West Point Workshop*, NY, USA, 2006, pp. 42–47.

[13]. V. Mura, D. Yambay, L. Ghiani, G. L. Marcialis, S. Schuckers, F. Roli, LivDet 2015 – Fingerprint Liveness Detection Competition 2015, in *Proceedings of the IEEE 7th International Conference on Biometrics: Technology, Applications and Systems (BTAS'15)*, 2015, pp. 1-6.

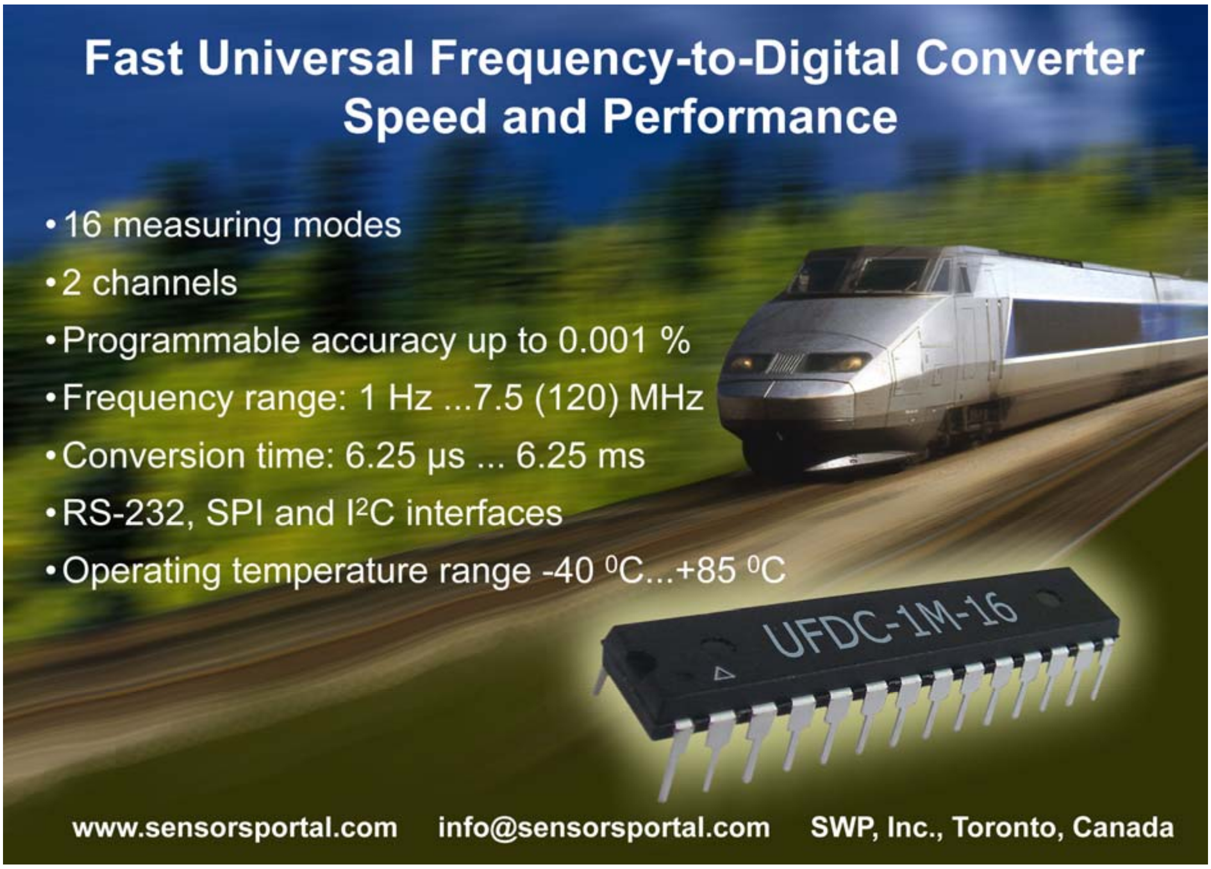
[14]. J. Galbally, J. Fierrez, R. Cappelli, An Introduction to Fingerprint Presentation Attack Detection, in Handbook of Biometric Anti-Spoofing. Advances in Computer Vision and Pattern Recognition (Marcel S., Nixon M., Fierrez J., Evans N., Eds), *Springer,* 2019, pp. 3-31.

[15]. G. Marcialis, F. Roli, A. Tidu, Analysis of fingerprint pores for vitality detection, *Proceedings of the 20th IEEE/IAPR International Conference on Pattern Recognition (ICPR'10)*, Instanbul, Turkey, 23-26 August 2010, pp. 1289–1292.

[16]. E. Marasco, C. Sansone, Combining perspiration- and morphology - based static features for fingerprint liveness detection, *Pattern Recognition Letters*, Vol. 33, Issue 9, 1 July 2012, pp. 1148-1156.

[17]. O. Czepulowski, S. Weidemann, A. Bailleu, Parallel Signal Acquisition by an Embedded System for Monitoring and Analysing Multimodal Signals of Aliveness and Non-aliveness of Biological Objects, in *Proceedings of the 11th International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies (BIOTECHNO'19)*, 02-06 June, 2019, Athen, Greece, 2019, pp. 8-12.

[18]. Microsystem for recognizing predeterminated features of valuable papers, security documents, or other products, *Patent EP 2 435 993 B1*, in 2016/27, 06.07.2016.



Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 8-15

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Seismic Noise and Site Response Analysis Using Accelerometer Sensors

*** Michele Paoletti, Roberto Concetti, Paola Pierleoni, Alberto Belli and Lorenzo Palma**
Università Politecnica delle Marche, 60131 Ancona, Italy
Tel.: 0712204847, fax: 0712204224
E-mail: m.paoletti@pm.univpm.it

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: We present an analyse of the environmental noise in absence of seismic activity detected by a seismic station (located in Gualdo in the province of Macerata equipped with two type of accelerometer sensors) and compare the average H/V noise curves with the results obtained using a velocimeter. The noise can be divided into various types: sensor noise, natural environmental noise, anthropic noise, microseismic noise. In this study we want to analyse the noise in frequency domain acquired by different sensors to understand if it is possible to extrapolate useful information from the seismic noise recording acquired by accelerometer rather than by velocimeter, in order to determine the sensor's limits.

Keywords: Seismic noise, Microtremor, Accelerometer, Ambient vibrations, H/V, HVSR.

1. Introduction

Seismic noise is any kind of seismic signal not generated by an earthquake and whose presence can alter the recording of a seismic event. The seismic noise (Fig. 1) is a stochastic stationary process rapidly varying over time: it can change from site to site and depends on many factors; for example, the noise generated by the wind is broadband and usually goes from 0.5 Hz to 15 Hz [1] while high-frequency sources are typically due to anthropic noise. In the presence of localized noise sources, the amplitude of the horizontal components is typically higher with respect to the vertical component. The seismic noise manifest as micro-vibrations that are acquired along the three axes by the sensor and that are visible if the acquisition system produce a noise lower than the seismic noise amplitude in the range of interest.

The statistical variability of the Power Spectral Density (PSD) is studied to identify the main causes of noise at the recording sites (assessing possible variations during the day) to determine the noise characteristic levels. In fact, studying the noise is also useful to obtain information on the detention thresholds of seismic events (in early warning systems the parameters are set, based on the noise, to detect better the seismic activity respect false positive). In microseismic noise evaluation, velocimeters are more sensitive than the accelerometers and therefore they are used to detect the fundamental frequency of a site. For this reason, in this analysis a velocimeter is used as reference.

A method for microzonation rather popular and convenient in terms of costs/benefits is based on the Nakamura technique that estimate the fundamental frequency evaluating the peak in the ratio of power spectral density of the horizontal component to the vertical component [2]. This technique is extremely advantageous because it allows the evaluation of the site response without the need to have a reference spectrum and the possibility of eliminating the effects of instrument responses.

Among seismic surveys, those based on the acquisition and analysis of environmental seismic noise or the continuous vibration of the soil due to both anthropic and natural causes, are becoming increasingly important for a rapid identification of the seismic substrate.

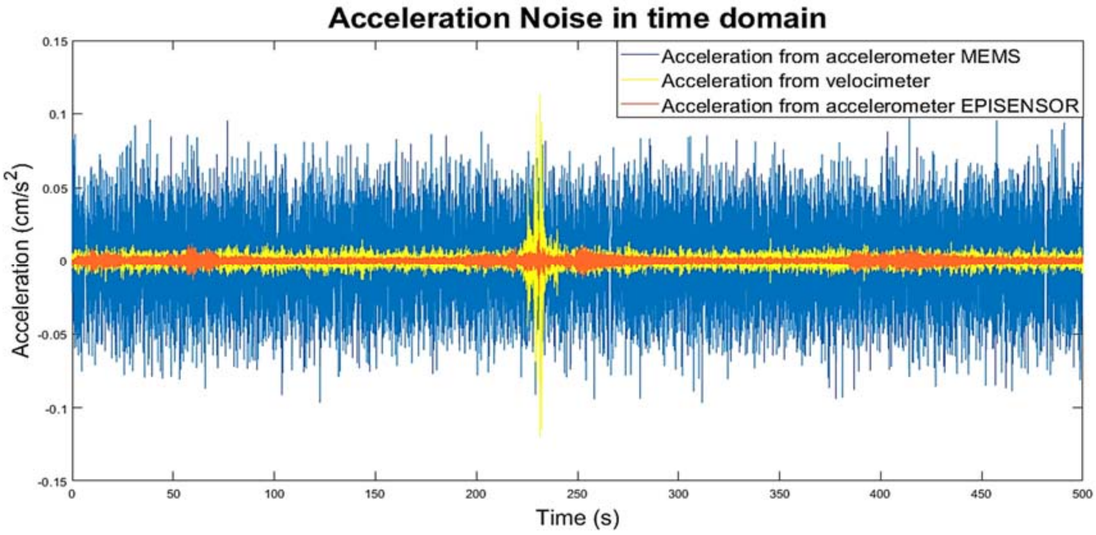


Fig. 1. Comparison between seismic noise accelerations (this is only one horizontal axis).

These types of techniques (defined as passive seismic methods) do not need any external energization because they use vehicle traffic, industrial production, wind, rain and all that can produce a minimal vibration on the source soil surface. Moreover, is totally non-invasive, very fast, applicable everywhere, low cost and not requires drilling. Using SSR (Standard Spectral Ratio) [3], it is possible to determine the amplification of the spectral components of the horizontal motion of a site with respect to a reference one; but this technique require a reference site where the site effects are absent (identify and separate the source from site effects is the main impediment encountered in spectral analysis). For these reasons other techniques should be investigated [4].

2. Method and Materials

The HVSR method (Horizontal to Vertical Spectral Ratios called also simply H/V), or Nakamura technique, is used to study the site response. When the HVSR technique is applied to seismic noise, often are used terms as HVSRN or NHVSR (Noise Horizontal Vertical Spectral Ratio), otherwise if it is applied to the seismic events EHVSR (Earthquake Horizontal Vertical Spectral Ratio) term is used. This passive HVSR method is known in literature as "Single Station" and it is defined as the ratio between the PSD spectral amplitudes of the horizontal components (mediating the two axes (1)) and the PSD spectral amplitudes of the vertical component of the motion [2, 5]. The HVSR technique is also useful to study the soil response during earthquake events and was applied for the first time to seismic recordings by [6] (taking only S-waves). This technique is based on the fundamental assumption that the vertical component of the ground motion is not influenced by the amplification effect caused by the trapping of seismic waves in the surface layer. Furthermore, the HVSR method can carry out measurements on an aseismic area using only seismic noise [4, 7].

The most important parameter, for a good NHVSR acquisition, is the recording time: the more you have a noisy environment (e.g. heavy road traffic in the vicinity, bad weather conditions, presence of industries etc.) and the longer the duration of the registration to be made.

The nature of the seismic noise wave is approximately the same as the propagation of surface waves [8-9] and a close relation exists between H/V spectral ratio and the ellipticity of the Rayleigh waves, which dominate the seismic noise.

The accelerometers used and compared in this analysis are:

- EpiSensor model FBA ES-T (external triaxial sensor) consists of three orthogonally mounted force balance low-noise accelerometers (FBAs) inside a sensor casing. EpiSensor is a triaxial accelerometer (X-axis, Y-axis and Z-axis) optimized for earthquake recording applications with high-performance, high-dynamic instrument (155 dB dynamic range), full g-scale range from ±0.25 to ±4 g (user selectable). The implemented technology allows to have very low self-noise, low non-linearity (<1000 μg/g^2), low hysteresis (<0.1 % of full scale), low cross-axis sensitivity (<1 %), and low thermal drift (<500 μg/°C) [12].

- Safran Colibrys model VS1002 capacitive MEMS accelerometer is used in order to measure seismic events in a three-dimensional space (3 Colibrys are positioned orthogonally to each other). The implemented technology allows to have a noise of $7\ g/\sqrt{Hz}$ in band, non-linearity < 0.1 % at full scale, full scale acceleration range of ±2 g, large and flat frequency response (±5 %) from DC to 700 Hz [12].

3. Signal Processing

The noise processing is made on data recorded from 1 March 2017 at 00:00:

- For the accelerometer a time span of nw=30 minutes (equal to 360.000 samples) and a window width of 100 s (equal to 20.000 samples) was used to calculate each H/V, obtaining n=18 windows with data acquired at a sample frequency fs = 200 Hz. Then the total average H/V is calculated using m=24 windows and a total time span equal to mw=12 hours (Fig. 2, Fig. 3, Fig. 4).
- For the velocimeter a time span of nw=60 minutes (equal to 360.000 samples) and a window width of 100 s (equal to 10.000 samples) was used to calculate each H/V, obtaining n=36 windows with data acquired at a sample frequency fs = 100 Hz. Then the total average H/V is calculated using m=24 windows and a total time span equal to mw=24 hours (Fig. 2, Fig. 5).

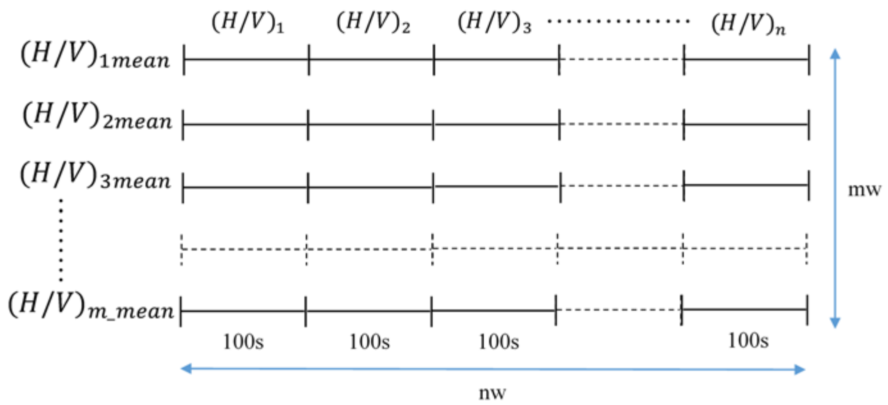


Fig. 2. Representation of H/V time intervals. H/V is calculated for each 100 s window. The mean is calculated for each row and then the total H/V average is calculated.

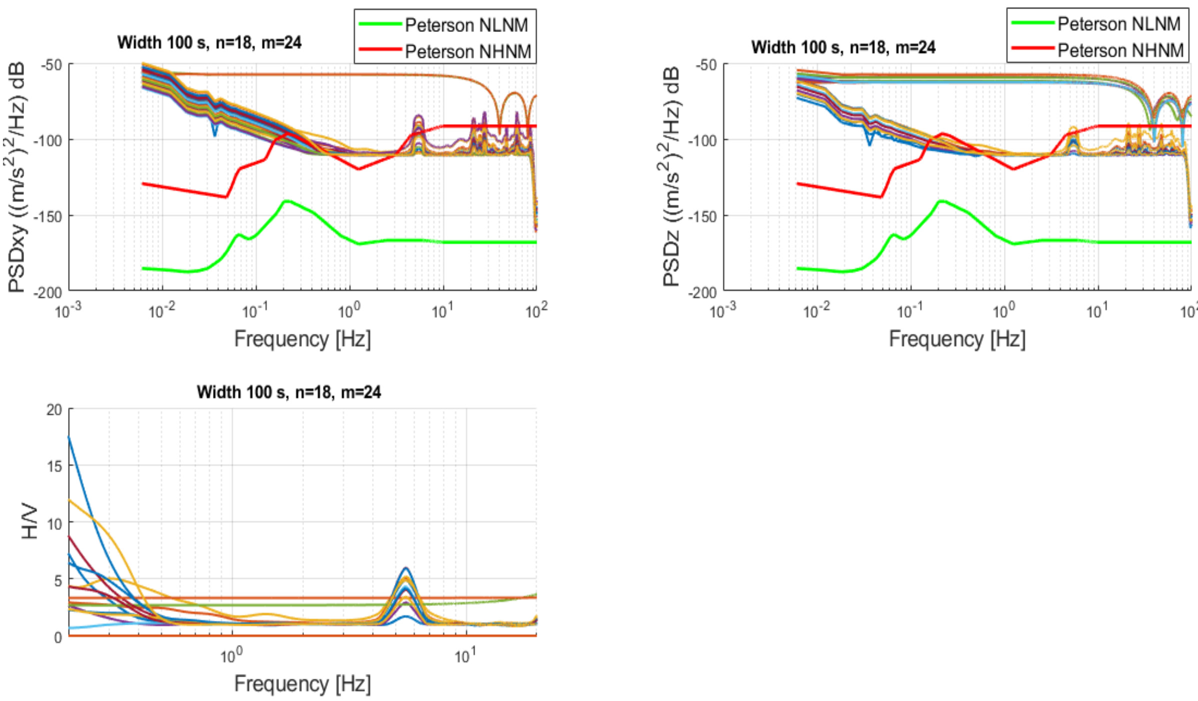


Fig. 3. Results using Safran Colibrys. PSDxy is the power density spectrum of horizontal components Hxy while PSDz of the vertical component.

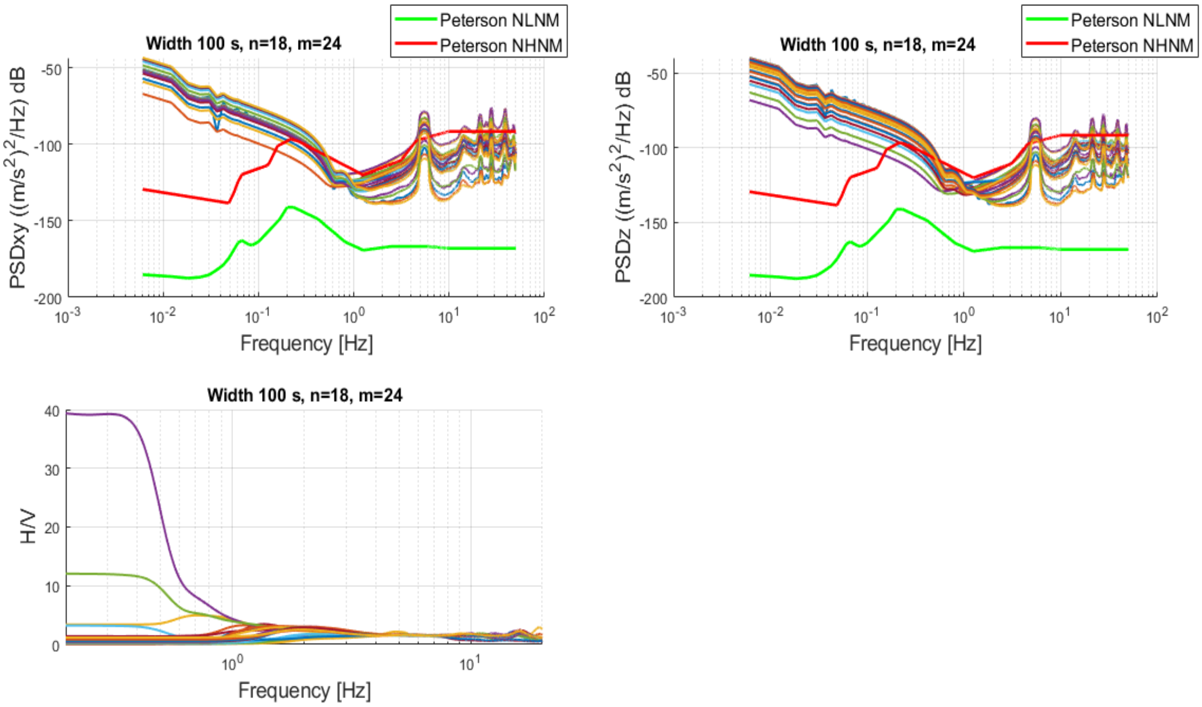


Fig. 4. Results using Episensor.

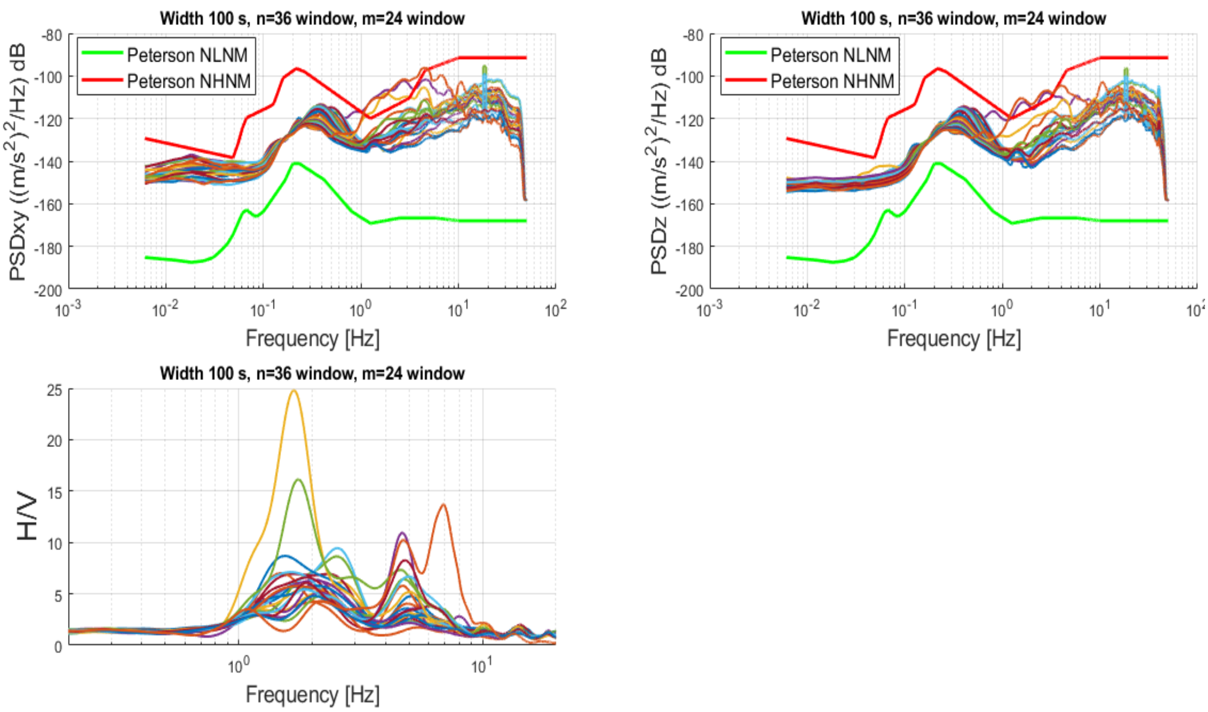


Fig. 5. Results using velocimeter.

In general, since the frequencies of engineering interest of the seismic and microseismic action do not exceed 25 Hz, the sampling frequency must not be less than 50 Hz (for the Nyquist-Shannon theorem). For the H/V technique it is also disadvantageous to use sample frequencies higher than 500 Hz, as they involve a considerable increase in the size of the files and consequently of processing times, without substantial improvements in the capacity or accuracy of the analysis. The duration of the acquisition window must be such as to characterize the natural frequency of interest. Therefore, if we have an idea of the first fundamental frequency value, then we can choose a window more appropriate; if instead this value is completely unknown, we will have to put ourselves in the most general conditions (that means to choose a greater window width to include the dynamics of the lower frequencies) [10].

As explained before in this analysis are used two type of accelerometers (with different features) and a velocimeter as reference to compare the results. In Fig. 1 shows, in time domain, the superimposed accelerations of one horizontal axis to see the amplitude difference (the acceleration of the velocimeter was obtained by deriving the velocity data). As expected, the noise is greater in MEMS accelerometer while for the Episensor and velocimeter the levels are similar (when no particular peaks of noise are present).

The programming language, used to implement the different elaborations, is Matlab. It allows to display the trace, the PSD, the PSD mean, the Peterson reference noise curves [11] and the spectral ratio between the horizontal and vertical component. Noise acquisitions are analyzed in absence of seismic activity to study the noise level of the environment and the site response.

These are the steps that characterize the noise analysis [4]:

1) The 3 components of the seismic noise are loaded, and the mean is removed.

2) A band pass filtering is carried out corresponding to the frequency range of interest (0.2-20 Hz). In the next graphs the filter is not applied to see the noise levels difference in a greater range.

3) The seismic noise is acquired using 100 s windows (Fig. 2.)

4) A 10 % cosine tapering is applied for each window to reduce the leakage phenomenon in frequency domain.

5) In each window the PSD is calculated for both the horizontal acceleration components and for the vertical component.

6) The 3 components are smoothed using Konno-Ohmachi smoothing function [13] (with parameter b = 40). It represents a "smoothing" of the amplitude spectrum to reduce stochastic variability in the estimation of spectral amplitude.

7) The quadratic mean is applied to the horizontal components [14-15] in order to obtain a single mean value (1):

$$H_{xy} = \sqrt{\frac{H_x^2+H_y^2}{2}}, \qquad (1)$$

where H_x means the East-West component and H_y the North-South component.

8) The H/V value is calculated for each window dividing the smoothed quadratic mean horizontal component PSD for smoothed mean PSD vertical component.

9) The H/V mean is determined by averaging the total number of windows (H/V mean or HVSR mean) spanning a total time of nw (Fig. 2).

10) The Konno-Ohmachi smoothing function is applied again to smooth out the obtained ratio with b=20 (to better reduce the variations).

11) The total average H/V is calculated spanning a total time of mw (Fig. 2).

A typical threshold value used in literature to consider a peak on the H/V curves is H/V=2.

4. Results

In Fig. 3, Fig. 4 and Fig. 5 are represented the analysis in frequency domain (the H/V curves are used then to obtain Fig. 6, Fig. 7, Fig. 8). In Fig. 3 is possible to see that until 3 Hz the noise in quite high compared to the NHNM Peterson curve. So, we expect that until 3 Hz is difficult to have any information about the site response because it is masked by the sensor noise. Between 3 Hz and 20 Hz the noise is quite normal, respect the Peterson curve, but with a bell around 5.5 Hz. Then, above 20 Hz the variations are outside the range of interest for site the response so we may not consider it. Are present also some very noisy curves that deviate sharply from the others and this could depend on some peaks of noise whose causes should be investigated.

In Fig. 4 the results are better, we expect that until 0.5 Hz is difficult to have information about the site response because they are masked by sensor noise (against the 3 Hz of the previous sensor). Also here at 5.5 Hz a bell is present in the PSD curves. In Fig. 5 the velocimeter results are better compared to the two previous cases.

In fact, there the noise is under the NHNM Peterson curve in the entire range and the differences between H/V curves are less.

5. Conclusions

From H/V of Fig. 3, Fig. 4 and Fig. 5 is possible calculate the average H/V where we can do the final consideration (Fig. 6, Fig. 7, Fig. 8). In Fig. 6 there is a small bell with a small peak at 5.5 Hz (the initial peak is not considered because the sensor is too unstable and not sensitive enough at very low frequencies, so the electronic noise produces large artificial amplitude in the H/V graphs).

In Fig. 7 there is a small peak at 2 Hz and another peak a 14.7 Hz (which is much smaller in the other curves) where is present a greater dispersion.

In Fig. 8 the curve is more reliable and average H/V have a peak about 1.8 Hz (the first resonance frequency) and 4.8 Hz (the second resonance frequency). The dispersion is greater between about 0.9 Hz and 9 Hz. The result obtained is very similar to the one provided by the INGV for the same station

(Fig. 9). In conclusion the accelerometers do not show sufficiently reliable and consistent results respect the velocimeter [16]. It would be interesting to repeat the same procedure on other days and other sites to obtain further confirmation. The other result from frequency analysis is that the Episensor accelerometer is better than Colibrys.

With the technology advancement and costs reduction more efficient accelerometer sensors will be produced and used to provide accurate data from a registration site together with the development of other different techniques to try to determine the site response. The use of the velocimeter remains of fundamental importance in the analysis of the seismic noise and site response.

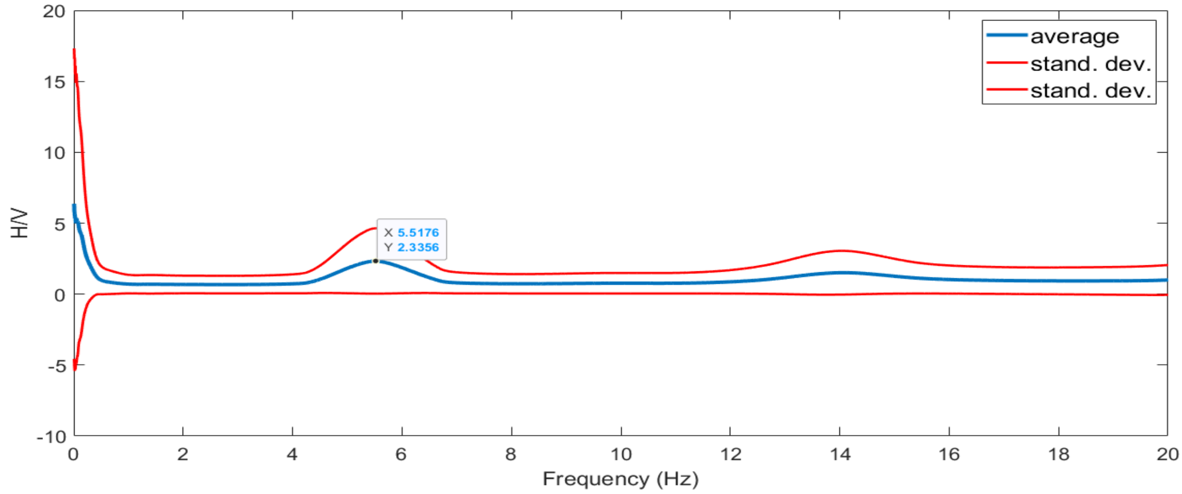


Fig. 6. Final results: H/V average and standard deviation of Colibrys accelerometer, total time span 12 hours.

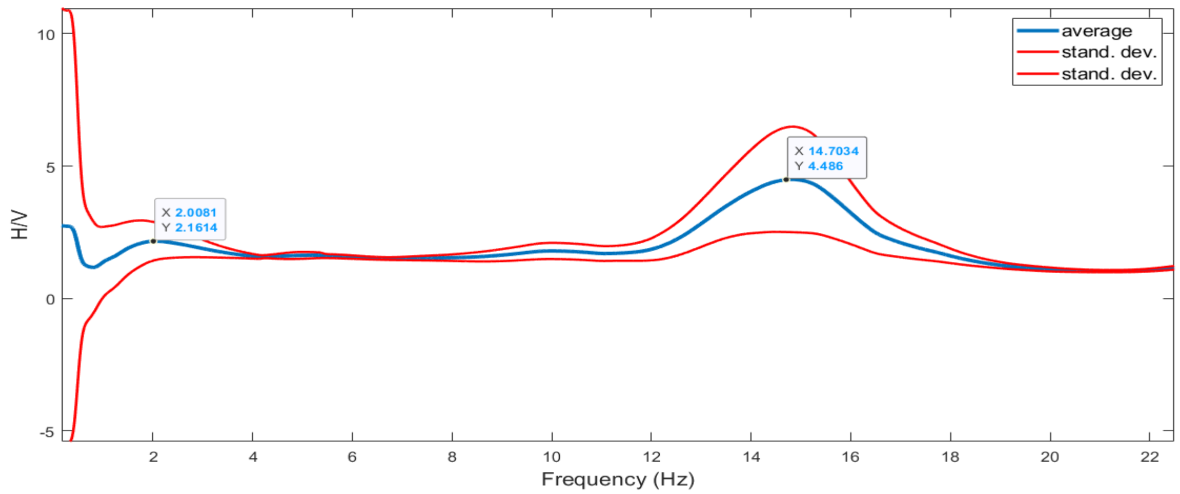


Fig. 7. Final results: H/V average and standard deviation of Episensor accelerometer, total time span 12 hours.

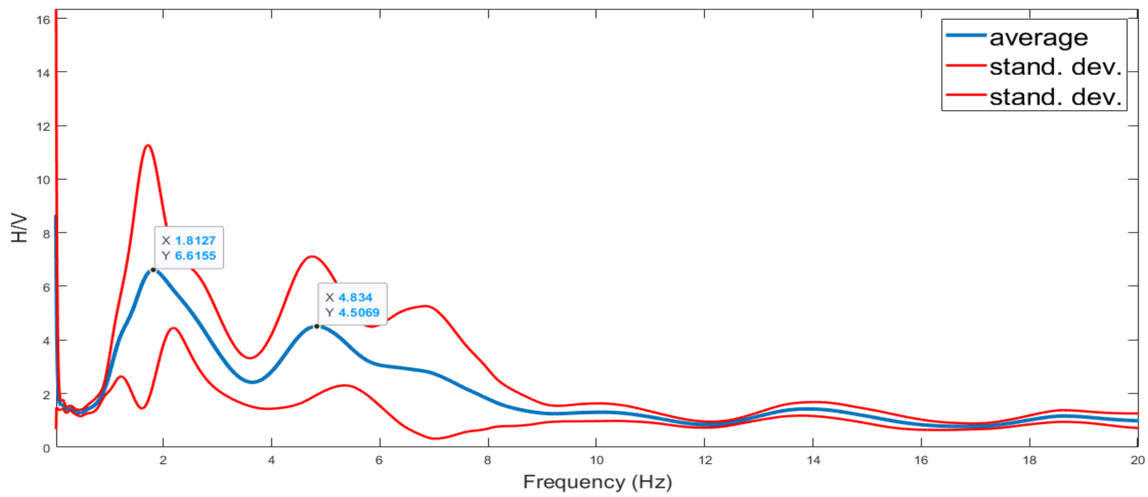


Fig. 8. Final results: H/V average and standard deviation of velocimeter, total span 24 hours.

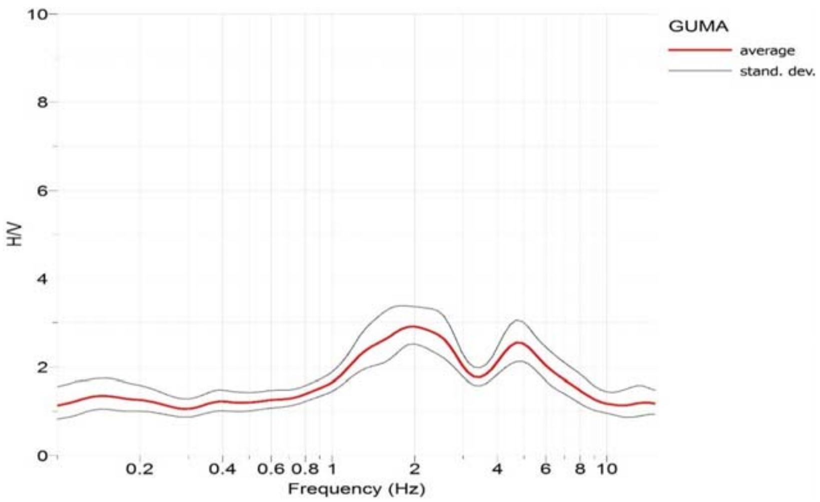


Fig. 9. Results obtained by INGV for the same station using a velocimeter but with different data and different software. This data is available in INGV website for GUMA station.

The knowledge of the resonance frequency of a soil, coupled to the information about the predominant period of a structure can give a reliable idea of potential damages that we can expect for a site in case of an earthquake. And this is of fundamental importance to reduce the effects that the earthquakes have on public and civil structures.

Acknowledgements

Thank you to the INGV for providing the raw data to be tested.

References

[1]. M. M. Withers, R. C. Aster, C. J. Young, E. P. Chael, High-frequency analysis of seismic background noise as a function of wind speed and shallow depth, *Bulletin of the Seismological Society of America*, Vol. 86, Issue 5, 1996, pp. 1507-1515.

[2]. Y. Nakamura, A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface, *Railway Technical Research Institute, Quarterly Reports*, Vol. 30, Issue 1, 1989, pp. 25-33.

[3]. R. D. Borcherdt, Effects of local geology on ground motion near San Francisco Bay, *Bulletin of the Seismological Society of America*, Vol. 60, Issue 1, 1970, pp. 29-61.

[4]. M. Paoletti, R. Concetti, P. Pierleoni, A. Belli, L. Palma, Spectral Analysis of Seismic Noise Using HVSR Technique, in *Proceedings of the 1st IFSA Frequency & Time Conference (IFTC'19), Tenerife, (Canary Islands), Spain, 25-27 September 2019, pp. 28-31*.

[5]. J. Lermo, F. J. Chavez-Garcia, Site effect evaluation using spectral ratios with only one station, *Bulletin of the seismological society of America*, Vol. 83, Issue 5, 1993, pp. 1574-1594.

[6]. H. Cruz-Jiménez, F. J. Chávez-García, T. Furumura, Differences in attenuation of ground motion perpendicular to the Mexican subduction zone between Colima and Guerrero: An explanation based on numerical modeling, *Bulletin of the Seismological Society of America*, Vol. 99, Issue 1, 2009, pp. 400-406.

[7]. M. Fazlavi, E. Haghshenas, Importance of mode detection in ambient noise array application for shear wave velocity profile determination, *International Journal of Civil Engineering*, Vol. 13, Issue 1, 2015, pp. 62-72.

[8]. K. Irikura, T. Kawanaka, Characteristics of microtremors on ground with discontinuous underground structure, *Bulletin of the Disaster Prevention Research Institute*, Vol. 30, Issue 3, 1980, pp. 81-96.

[9]. P. Guéguen, C. Cornou, S. Garambois, J. Banton, On the limitation of the H/V spectral ratio using seismic noise as an exploration tool: application to the Grenoble valley (France), a small apex ratio basin, *Pure and Applied Geophysics*, Vol. 164, Isuue 1, 2007, pp. 115-134.

[10]. E. Haghshenas, P. Y. Bard, N. Theodulidis, Sesame WP04 Team, Empirical evaluation of microtremor H/V spectral ratio, *Bulletin of Earthquake Engineering*, Vol. 6, Issue 1, 2008, pp. 75-108.

[11]. J. R. Peterson, Observations and modeling of seismic background noise, *US Geological Survey*, 1993.

[12]. P. Pierleoni, S. Marzorati, C. Ladina, S. Raggiunto, A. Belli, L. Palma, M. Cattaneo, S. Valenti, Performance Evaluation of a Low-Cost Sensing Unit for Seismic Applications: Field Testing During Seismic Events of 2016-2017 in Central Italy, *IEEE Sensors Journal*, Vol 18, Issue 16, 2018, pp. 6644-6659.

[13]. K. Konno, T. Ohmachi, Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bulletin of the Seismological Society of America*, Vol. 88, Issue 1, 1998, pp. 228-241.

[14]. D. Albarello, E. Lunedei, Combining horizontal ambient vibration components for H/V spectral ratio estimates, *Geophysical Journal International*, Vol. 194, Issue 2, 2013, pp. 936-951.

[15]. B. C. Sylvette, C. Cécile, B. Pierre-Yves, C. Fabrice, M. Peter, K. Jozef, D. Fäh, H/V ratio: a tool for site effects evaluation. Results from 1-D noise simulations, *Geophysical Journal International*, Vol. 167, Issue 2, 2006, pp. 827-837.

[16]. M. Mucciarelli, Reliability and applicability of Nakamura's technique using microtremors: an experimental approach, *Journal of Earthquake Engineering*, Vol. 2, Issue 4, 1998, pp. 625-638.

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 16-20

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Scanning Electron Microscopy Imaging of Acoustic Wave Fields Distribution in QCM

[1, *] Dmitry ROSHCHUPKIN, [1] Evgenii EMELIN, [1] Olga PLOTITCYNA and [2] Sergei SAKHAROV
[1] Institute of Microelectronics Technology and High Purity Materials Russian Academy of Sciences, Chernogolovka, Moscow Region, 142432, Russia
[2] FOMOS Materials Co., Buzheninova St. 16, Moscow, 105023, Russia
[1] Tel.: +74965244058, fax: +74965244225
[*] E-mail: rochtch@iptm.ru

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: The acoustic properties of QCM microbalances based on a $La_3Ga_5SiO_{14}$ crystal (LGS) were studied using scanning electron microscopy technique. Scanning electron microscopy method in the low-energy secondary electron emission mode makes it possible to visualize the distribution of the acoustic wave field in the QCM (distribution of nodes and antinodes of standing bulk acoustic wave in crystal) in the real time mode.

Keywords: $La_3Ga_5SiO_{14}$ crystal, Scanning electron microscopy, Bulk acoustic waves, QCM, Microbalance.

1. Introduction

Usually, electrical measuring methods are used to analyze the functioning of acoustoelectronic devices based on surface and bulk, traveling and standing acoustic waves. It should be noted that these methods do not allow analyzing the process of propagation of acoustic waves in solids. For the analysis of acoustic fields in solids, it is advisable to use the methods of scanning electron microscopy (SEM) [1-3] and X-ray topography [4-6], which allows to visualize the acoustic wave fields distribution both on the surface and in the volume of crystals. These methods make it possible to measure the acoustic wavelengths, to determine the direction of propagation of the acoustic signal and power flow angles, to investigate the acoustic wave fields interaction with crystal structure defects.

The purpose of this work was to study the distribution of the acoustic wave field in QCM microbalance using the scanning electron microscopy method.

2. Scanning Electron Microscopy Imaging of Acoustic Wave Fields: Experimental Set-Up

The SEM method in the mode of secondary electron recording was used as an experimental technique for visualization of the bulk acoustic waves excitation and propagation in QCM microbalance since only the low energy secondary electrons with energy of ~1÷3 eV are sensitive to the electric field which accompanies the propagation of acoustic waves in piezoelectric materials. The acoustic waves were rendered visible in a JEOL JSM-840 scanning electron microscope with an accelerating voltage E_0=1÷3 kV and a probe current $I_0=6\times10^{-9}$ A. The use of high accelerating voltage is not possible, because the insulating substrate would be highly charged and thus bring about a distortion of the image because of the deflection of the primary electron beam and the change in the secondary electron emission from the crystal surface.

For visualization of acoustic wave propagation in solids, it is possible to use two different modes of SEM method. In the first one, it is possible to use the SEM in the stroboscopic mode with high-frequency modulation of the electron beam with the frequency of acoustic wave. Under this condition, the electron beam is in phase synchronization with acoustic wav and secondary electron emission from the surface is defined by the electric filed, which accompanies the propagation of acoustic wave in piezoelectric materials [1]. In the second case, which is more handy (it is not necessary to use the special stroboscopic techniques), the high-frequency modulation of the low energy secondary electrons is defined by a stationary electrostatic interference field formed above the crystal surface in the process of interaction between the varying electrostatic field of acoustic wave and the component, normal to the crystal surface, of the electromagnetic radiation field of QCM-electrodes. The electromagnetic and acoustic waves are mutually coherent, since they are excited by the same source (QCM-electrodes) and with the same frequency. Since the wavelength of electromagnetic wave is large than the wavelength of acoustic wave (~five order of magnitude), the period of the stationary interference field is equal to the wavelength of acoustic wave. Under these conditions, the low energy secondary electrons are modulated by the stationary electrostatic interference field. In our investigations to visualize the acoustic waves in QCM we used the second method [2-3].

SEM method was used to study of the process of bulk acoustic wave excitation in QCM based on *yxl*/+50°-cut of a $La_3Ga_5SiO_{14}$ crystal (Fig. 1).

Fig. 1. QCM microbalances.

Different types of electrode materials (Au and Ir) were used. Usually bulk acoustic waves (BAW) are excited in QCM not only at resonance excitation frequency, but also at some another frequencies due to accuracy of QCM fabrication, QCM surface curvature accuracy, location and structure of metal electrodes on the surface of QCM.

3. SEM Imaging of Acoustic Wave Fields in QCM Microbalances

Fig. 2 shows the amplitude frequency response (AFR) S11 of QCM with Au-electrode based on transverse BAW. It seen that resonance excitation frequency is f_0=6.0072 MHz, while the frequency of antiresonance is f=6.0150 MHz. Fig. 2 also display the different excitation frequencies due to the precision of QCM fabrication. At f=6.1210 MHz QCM excited the transverse BAW, that propagates along axis x.

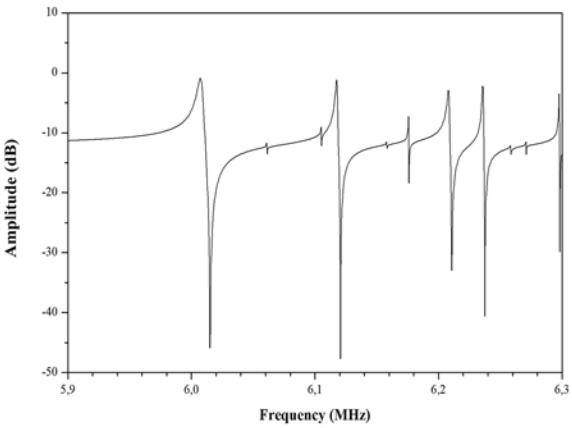
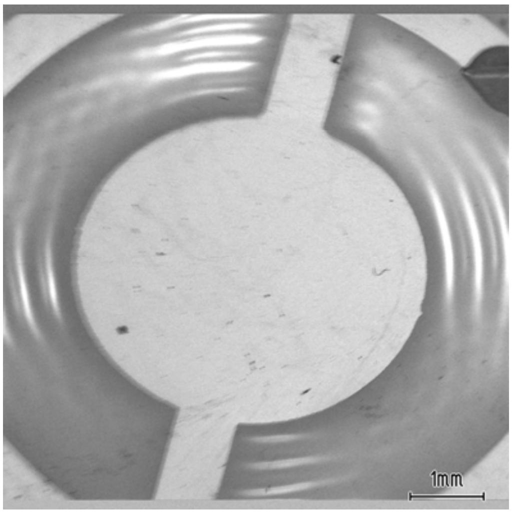


Fig. 2. AFR S11 of QCM with Au-electrodes.

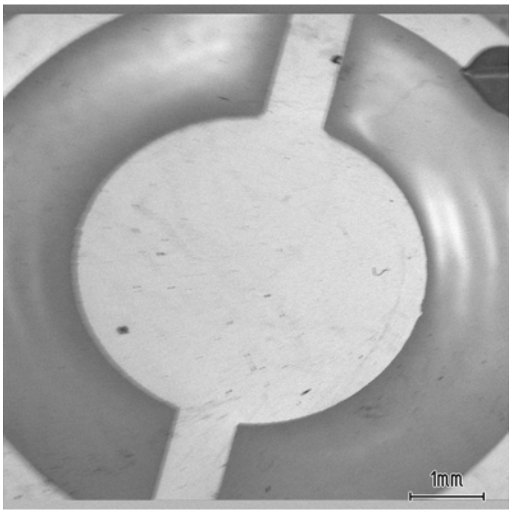
Fig. 3 demonstrated the SEM images of BAW excited in QCM microbalance at different excitation frequencies. At f_0=6.0072 MHz Fig. 3(a) shows the resonance excitation mode of transverse BAW, that propagates along axis x. Fig. 3(b) shows the BAW distribution in QCM at antiresonace conditions of f=6.0150 MHz. The distribution of BAW in nonresonance mode at f=6.1210 MHz is shown in Fig. 3(c). Thus, the SEM method is very simple method to study the operation of QCM based on bulk acoustic wave.

Fig. 4 shows the amplitude frequency response S11 of QCM with Ir-electrode based on transverse BAW. It seen that resonance excitation frequency is f_0=5.9980 MHz. AFR at Fig. 4 also demonstrates the different excitation frequencies due to the precision of QCM fabrication (undesirable parasitic modes). At f=5.9980 MHz QCM excited the transverse BAW that propagating along axis x.

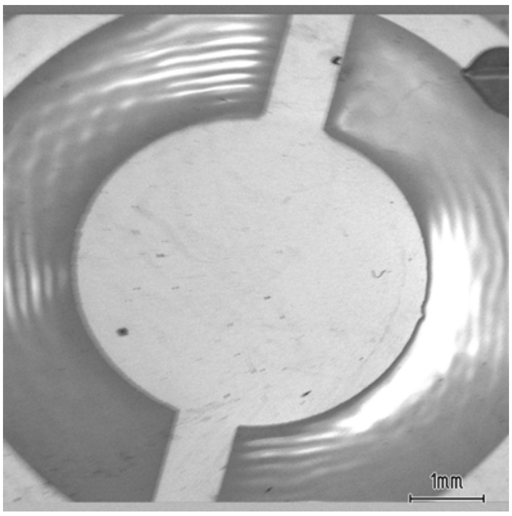
Fig. 5 shows the SEM images of BAW excited in QCM microbalance with Ir-electrodes at different excitation frequencies. At f_0=5.9980 MHz Fig. 5(a) display the resonance excitation mode of transverse BAW, that propagates along axis x. Fig. 5(b) shows the BAW in nonresonance mode at f=6.1045 MHz.



(a)



(b)



(c)

Fig. 3. SEM images of BAW in QCM with Au-electrodes excited at resonance excitation frequency f_0=6.0072 MHz (a), f=6.0150 MHz (b), f=6.1210 MHz (c).

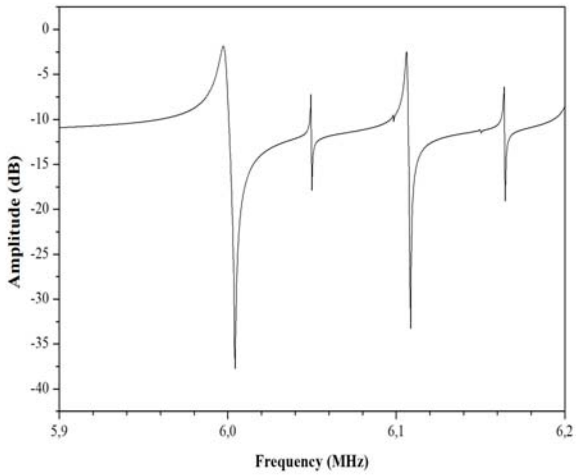
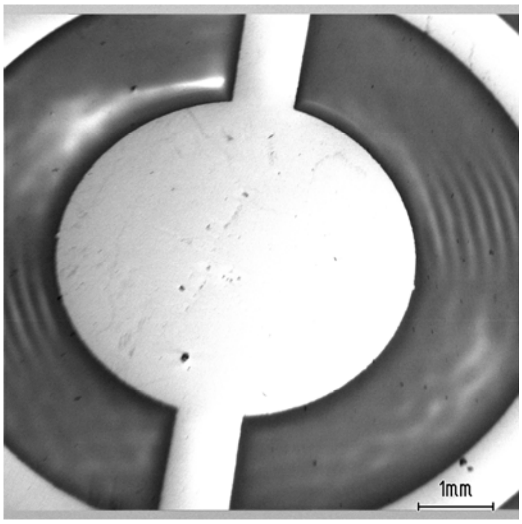
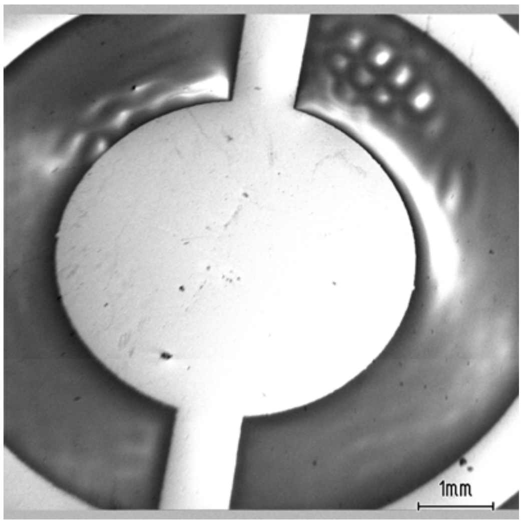


Fig. 4. AFR S11 of QCM with Ir-electrodes.



(a)



(b)

Fig. 5. SEM images of BAW in QCM with Ir-electrodes excited at resonance excitation frequency f_0=5.9980 MHz (a), f=6.1045 MHz (b).

Distribution of acoustic wave field in QCM strongly depends on crystal and electrodes forms while the resonance excitation frequency depends only from crystal thickness.

Fig. 6 shows the result of SEM investigation of QCM operated at more high frequency. Fig. 6(a) demonstrates the distribution of resonance excitation mode of transverse bulk acoustic wave in QCM at resonance excitation frequency of f_0=22.0247 MHz. Figs. 6(b)-(c) demonstrates the parasitic modes of transverse acoustic waves in QCM excited at *f*=22.1064 MHz and *f*=22.1529 MHz, respectively. In this case the transverse acoustic waves propagates along axis x.

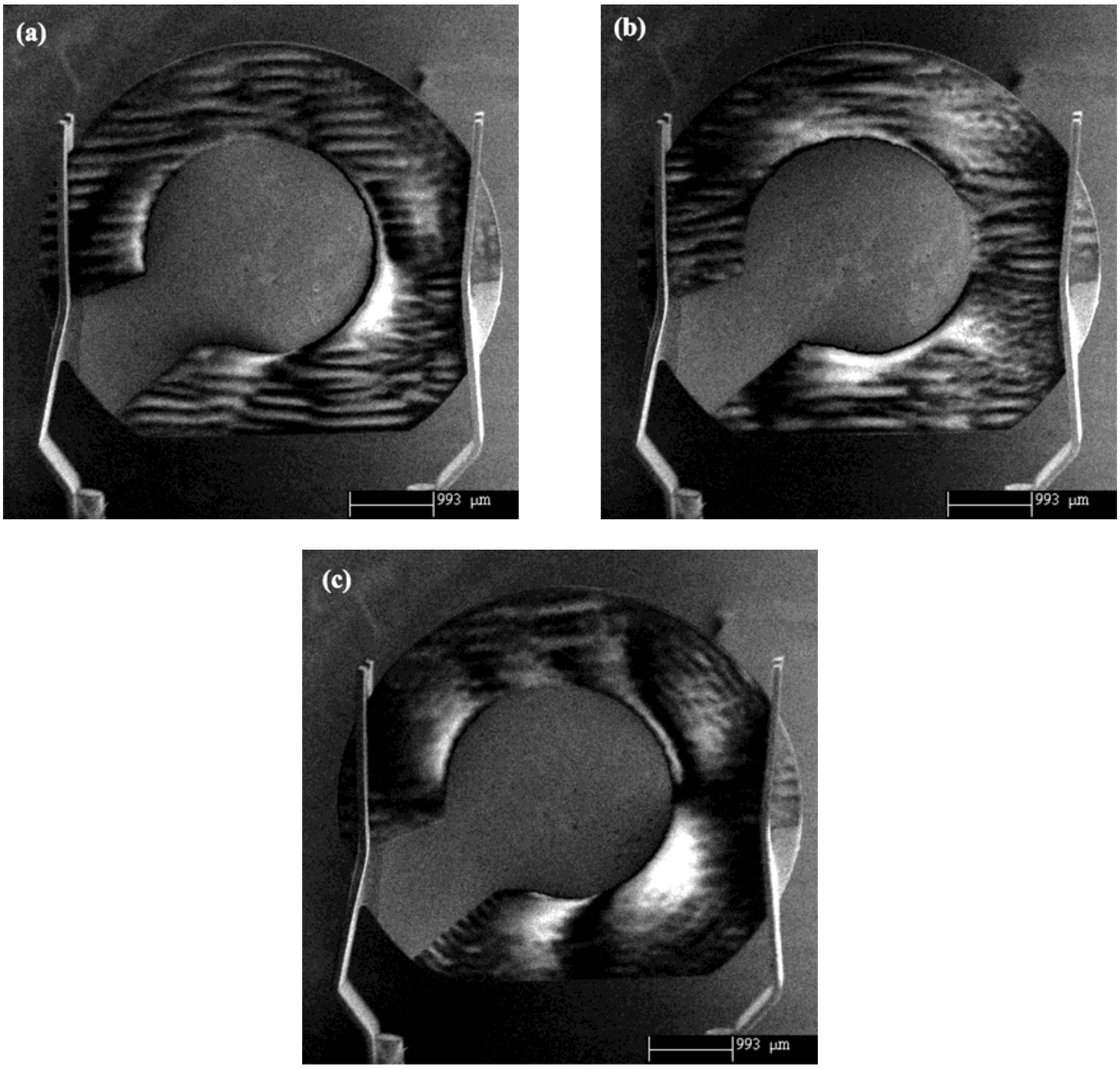


Fig. 6. QCM based on *yxl*/+50°-cut of a $La_3Ga_5SiO_{14}$ crystal: (a) resonance excitation frequency f_0=22.0247 MHz, (b) *f*=22.1064 MHz, (c) *f*=22.1529 MHz.

4. Conclusions

We demonstrate the real possibility to use the scanning electron microscopy method for visualization and controlling of acoustic wave fields in QCM microbalance. It is shown, that SEM method is very fast and useful method for investigation of acoustoelectronic devices.

The only drawback of the SEM method is that it is impossible to observe the image of acoustic waves under the metal electrode, which is an equipotential surface.

References

[1]. H. P. Feuerbaum, G. Eberharter, G. Tobolka, Visualization of traveling surface acoustic waves using a SEM, *Scanning Electron Nicroscopy*, Vol. 1, 1980, pp. 502-508.

[2]. D. V. Roshchupkin, M. Brunel, Scanning electron microscopy observation of surface acoustic wave propagation in the $LiNbO_3$ crystals with regular domain structures, *IEEE Transaction on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 41, Issue 4, 1994, pp. 512-517.

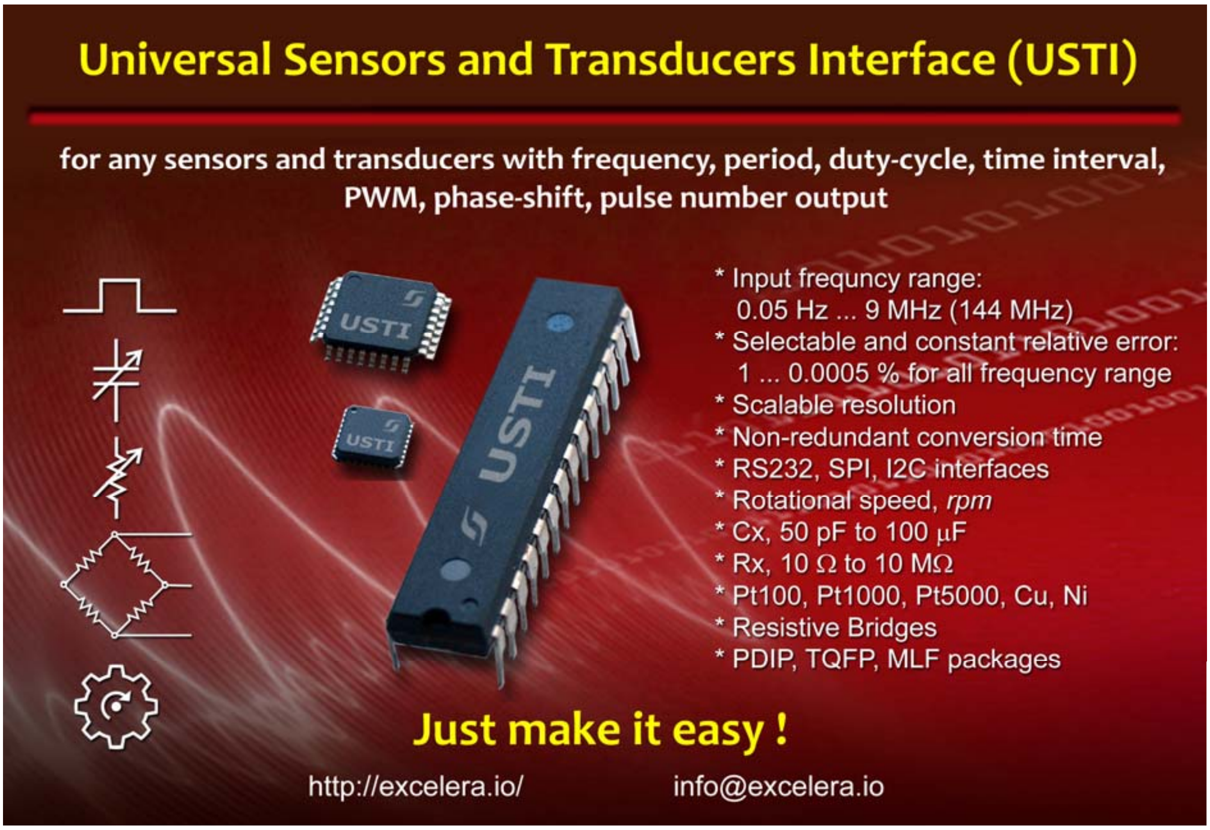
[3]. D. V. Roshchupkin, Th. Fournier, M. Brunel, O. Plotitsyna, N. G. Sorokin, Scanning electron

microscopy observation of excitation of the surface acoustic waves by the regular domain structures in the $LiNbO_3$ crystals, *Appl. Phys. Lett.*, Vol. 60, Issue 19, 1992, pp. 2330-2331.

[4]. D. V. Roshchupkin, H. D. Roshchupkina, D. V. Irzhak, X-Ray topography analysis of acoustic wave fields in the SAW-resonators structures, *IEEE Transaction on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 52, Issue 11, 2005, pp. 2081-2087.

[5]. D. Roshchupkin, L. Ortega, A. Snigirev, I. Snigireva, X-ray imaging of the surface acoustic wave propagation in $La_3Ga_5SiO_{14}$ crystal, *Appl. Phys. Lett.*, Vol. 103, Issue 15, 2013, pp. 154101-1–154101-3.

[6]. D. Roshchupkin, L. Ortega, I. Zizak, O. Plotitcyna, V. Matveev, O. Kononenko, E. Emelin, A. Erko, K. Tynyshtykbayev, D. Irzhak, Z. Insepov, Surface acoustic wave propagation in graphene film, *Journal of Applied Physics*, Vol. 118, Issue 10, 2015, pp. 104901-1–104901-5.



Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 21-30

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

A Low Cost Ultrasound-based Localisation System for Ground Robotics

Alec BURNS, Sebastiano FICHERA and * Paolo PAOLETTI
School of Engineering, University of Liverpool, Brownlow Hill, Liverpool L69 3GH, UK
Tel.: 44 151 7945232
* E-mail: P.Paoletti@liverpool.ac.uk

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: This paper presents a low-cost localisation system based on ultrasonic sensing and time of flight measurements. A compact ultrasound emitter has been designed to generate an omnidirectional train of ultrasound pulses which are then picked up by several fixed receivers measuring the time difference of arrival. A least squares approach is used to analytically obtain a first estimate of the emitter position, which is then refined through steepest descent optimisation. All processing is done via a standard Arduino platform, proving the low computational demands of the method. Localisation results are validated against a state-of-the-art Optitrack motion capture system. It is shown that the system can cover a 4.3×3.1 m arena with a mean error localisation error of 1.57 cm and an average standard deviation of 1.39 cm throughout the arena. The effectiveness of the proposed localization system is demonstrated by integrating it with a mobile ground robot to enable waypoint-based path following.

Keywords: Localisation, Ultrasound, Time difference of arrival, Positioning, Ground robots.

1. Introduction

Ground robots are becoming increasingly popular platforms for research, industry and hobbyists alike. As the rise in demand has continued, so too has the need for these devices to be compact, simple and inexpensive, yet accurate and robust. The ability to determine their position within the surrounding environment is a critical enabling capability in order for a mobile robot to move around and interact with large scale environments [1].

Over the years, several methods have been developed to tackle the localization challenge. For lower-end robots, i.e. simple low-cost robots with limited computational power, dead-reckoning methods have been widely used. This method is based on local sensors feedback, e.g. encoder readings in wheeled robots. Knowledge of the starting position and distance travelled allows one to estimate current position. However, small inaccuracies and drifts in the system, such as slippage of the wheels, may lead to large errors over long distances if not properly compensated for [2].

Another popular option is Visual Positioning Systems (VPS) that compare the current scene with previously stored scenes to create a best estimation of current location. Recent developments in the field have made this technique a viable option for localisation [3]. However, the relatively high computational cost and hardware complexity make this approach not suitable for implementing a real-time low-cost localisation system [4]. Also, VPS require a prior knowledge of the map within which they will operate, preventing them to be used in unknown, dynamic or unstructured environments.

Localisation solutions based on external sensory networks may provide an effective solution for satisfying accuracy requirements while meeting low-cost constraints. One such solution is the Active Bat system, a broadband ultrasonic position system utilizing Time-of-Flight information from the badges to a network of receivers on the ceiling and calculates the 3D positions of the badges using a multilateration algorithm. This system achieves a 2D accuracy of approximately 2 cm. However, the coverage of 1000 m^2 is attained with 750 receivers required [5].

Another example is Lok8, an indoor positing system for smartphones that uses off-the-shelf smartphone speakers to produce ultrasound signals (operating at around 22 kHz). The systems accuracy was tested in a 49 m^2 with 4 receivers and achieved an accuracy of around 10cm on average [6].

Commercially available systems can achieve precision up to ± 2 cm while costing less then £500, see [7] for an example which relies on a series of stationary ultrasonic receiver modules and one mobile emitter module. The location of the mobile emitter module is calculated based on propagation delay of the signal. Another commercially available localisation solution, which utilizes a similar positioning algorithm is described in [8]. This product relies on Ultra-Wide Band (UWB) wave propagation instead of ultrasonic signals, is available for just under $1000 and is a centimetre-level accurate positioning system.

In this work, a novel low-cost localisation system based on ultrasound sensing is described and validated against a state-of-the-art motion tracking system. It is shown that such localisation platform is able to provide good accuracy while requiring very limited sensing and computational complexity.

This paper is organised as follows. Section 2 motivates the design choices for the localisation system, which is then presented in detail in Section 3. Section 4 describes the hardware setup used to assess the quality of the proposed solution, and results from such assessment are discussed in Section 5. Finally, some conclusions are drawn in Section 6, together with suggestions for future research directions.

2. Definition of the Infrastructure

Various factors need to be taken in consideration when designing or assessing any localisation technique, including: accuracy, precision, robustness, financial cost and system simplicity (refer to [9] and [10] for definitions of these metrics).

Wireless indoor localisation approaches can be classified according to two main criteria [9-10].

- Physical sensor infrastructure, i.e. the platform used to detect/sense position;
- Positioning algorithm, i.e. the method to estimate location from sensory data.

The combination of these elements determines the overall method for localising the mobile robots.

2.1. Sensor Topology

Most physical sensory infrastructures rely on two aspects: a signal emitter and a measuring unit called a receiver. Measurement involves the transmission and reception of signals between these parts of the system. There are four different system topologies for positioning systems [9]. The Remote Positioning System (RPS) places the emitter unit on the mobile robot at an unknown location. The emitter generates a signal which is received by receivers placed at known fixed locations. The results from the measuring units are collected and the location of the emission source calculated on an external master station. The second topology is Self-Positioning System (SPS) and involves a mobile unit which receives signals of several transmitters placed at known locations and has the capability to compute its location based on the measured signals. Sending the final measurement results from the mobile unit to the remote unit in an SPS is referred to as indirect self-positioning system (ISPS), which is the third type of topology. Finally, transmitting the final measurement result from the remote unit to the mobile unit in an RPS is referred to as Indirect Remote Positioning System (IRPS).

IRPS was selected here as it reduces the need for a mobile robot to have high computational capabilities on board and offsets this function to a ground station. This widens the range of systems the localisation platform could be applied to, especially when considering simple low-level robots, robots with small payload or swarm systems. Furthermore, transmitting the final results from the remote unit to the mobile unit allows all decision making to be done on the mobile robot and, potentially, enables sensor fusion with odometry.

2.2. Sensor Types

Within the IRPS family of localisation systems, different choices of signals propagated between the emitter and receivers can be made. Table 1 provides a comparison of the most popular methods available.

The typical dimension of low-cost mobile ground robots is of the order of tens of centimetres, therefore it is reasonable to assume that the localisation system needs to have an accuracy of less than 5 cm. As can be seen from Table 1, this requirement rules out solutions such as RFID, ZigBee and WiFi. UWB could be an option for future development as it is undergoing fast development in terms of accuracy, has strong anti-interference capability and is capable of permeating through objects [12]. However, state of the art UWB-based systems are still not accurate enough to meet the requirement outlined above. For example, in [13] authors achieved accuracies <31 cm in 90 % of position estimates with an average inter-sample noise of 9 cm.

Ultrasound (US) localisation was therefore selected as the most appropriate sensor type due to its potential high accuracy. Furthermore, due to the slow propagation speed of US waves (340 m/s), simple processing technology can be used, reducing overall complexity. The only drawback of current systems are overall cost and scalability [4, 14]. The system proposed here overcomes these issues, maintaining high accuracy on a large arena while being significantly less expensive than commercially available platforms.

The use of the ultrasonic platform within an IRPS topology is novel due to the difficulties presented in designing a central 'omnidirectional' ultrasonic emitter, the details of which are presented in Section 4.1.

Table 1. Comparison of main indoor positioning technologies [4, 10].

System	Accuracy	Coverage	Methods	Topology	Scale	Complexity	Cost
IR	57 cm-2.3 m	Room	Cell-ID, RSSI, TOA	(I)SPS, (I)RPS	Low	High	High
UWB	15 cm	Building	RSSI, TOA, TDOA, AoA	(I)RPS	Low	Low	High
US	1 cm-2 m	Room	ToA, TDoA	(I)RPS	Low	Low	High
WiFi	1.5 m	Building	Cell-ID, RSSI, TOA, AoA	(I)SPS, (I)RPS	High	Low	Low
BT	30 cm-meters	Building	Cell-ID, RSSI, TOA	(I)SPS	High	Low	High
RFID	1-5 m	Room	Cell-ID, RSSI	(I)SPS,(I)RPS	Medium	Low	Low
ZigBee	30 cm-meters	Building	Cell-ID, RSSI	(I)SPS, (I)RPS	Low	Low	Medium
Audible Sound	Meters	Room	ToA, TDoA	(I)RPS	Low	Low	High
GPS	6 m-10 m	Global	ToA	(I)SPS	Low	High	High

3. Localisation Estimation Algorithm

The proposed system consists of a central 'omnidirectional' point emitter to be localised and a series of fixed directional receivers at known locations.

Time-Difference-of-Arrival (TDoA) was selected as the approach to estimate emitter position. When compared with other approaches such as Angle-of-Arrival or Time-of-Arrival, TDoA systems are cheaper and simpler in both hardware and computational algorithm [4, 15, 16]. For TDoA, an omnidirectional US signal is generated from the mobile emitter. As this pulse reaches the first receiver, the time of reception t_0 is recorded. Furthermore, as the US pulse continues to travel, it will continue triggering receivers which will also record times *relative* to the first receiver $t_{i0} = t_i - t_0$. These relative times of reception allow localization of the emitter on a hyperboloid with the first receiver and i-th receiver positions as foci [17]. Mathematically, this translates into the system of equations [18].

$$\begin{aligned}(x-x_0)^2+(y-y_0)^2+(z-z_0)^2&=(t_0+t_{00})^2v_s^2\\(x-x_1)^2+(y-y_1)^2+(z-z_1)^2&=(t_0+t_{10})^2v_s^2\\&\vdots\\(x-x_N)^2+(y-y_N)^2+(z-z_N)^2&=(t_0+t_{N0})^2v_s^2\end{aligned},\qquad(1)$$

where v_s is the speed of the US signal, (x,y,z) are the unknown emitter coordinates and (x_i,y_i,z_i) is the known positions of the i-th receiver. Note that $t_{00} = 0$ by definition.

3.1. Initial Localisation via Triangulation

The physical infrastructure described in Section 2 and the TDoA approach described in Section 3 provides a series of hyperboloids theoretically intersecting at the emitter location. However, noise in signal propagation, received signal measurement and location of receivers translates to an uncertainty in the estimation of emitter position. Closed-form solutions for the localisation problem do not accommodate for situations where hyperboloids do not intersect at a single point [19], therefore best approximations algorithms are always necessary. An analytical solution to best-fit the emitter position in TDoA systems has been proposed in [18] and is used as a first step in the algorithm proposed in this paper.

To this end, the first line of Eq. (1) is subtracted from the subsequent lines, thus obtaining

$$2A\boldsymbol{p}=\boldsymbol{b},\qquad(2)$$

where

$$A=\begin{bmatrix}x_1'\,y_1'\,z_1'\,t_{10}v_s^2\\x_2'\,y_2'\,z_2'\,t_{20}v_s^2\\\vdots\\x_N'\,y_N'\,z_N'\,t_{N0}v_s^2\end{bmatrix},\ \boldsymbol{p}=\begin{bmatrix}x\\y\\z\\t_0\end{bmatrix}$$

$$\boldsymbol{b}=\begin{bmatrix}x_1^2-x_0^2+y_1^2-y_0^2+z_1^2-z_0^2-t_{10}v_s^2\\x_2^2-x_0^2+y_2^2-y_0^2+z_2^2-z_0^2-t_{20}v_s^2\\\vdots\\x_N^2-x_0^2+y_N^2-y_0^2+z_N^2-z_0^2-t_{N0}v_s^2\end{bmatrix},\qquad(3)$$

and $x_i' = x_i - x_0, y_i' = y_i - y_0, z_i' = z_i - z_0, i = 1, \dots, N$.

The solution of this system of linear equations then provides the emitter position $\boldsymbol{p}$. As mentioned before, presence of noise and uncertainties implies that no exact solution exists, therefore a Least Squares algorithm was used here to solve Eq. (1) with minimal computational cost [20]

$$\boldsymbol{p} = \left(\frac{1}{2}\right)(A^T A)^{-1} A^T(\boldsymbol{b}) \tag{4}$$

Note that this solution is simpler than the one described in [18] as it provides $\boldsymbol{p}$ in a single step and does not involve second order equations which may lead to multiple solutions. However, as matrix inversion $(A^T A)$ is required, the matrix $A^T A$ must be non-singular. According to the definition of the matrix A in Eq. (3), singular configurations may occur if, for example, all the receivers lay on the same plane (in this case one of the first three columns is zero), or if more than *N-4* receivers are in the same point (in this case *N-4* rows are linearly dependent). These pathological situations can be easily avoided by placing the receivers accordingly. Lastly, errors will occur if all values of the fourth column of matrix A are *exactly* the same, that is if all receivers are the exact same distance away from the closest receiver to the robot. Once again, this situation is unlikely to occur in practice if the receivers are positioned correctly. Therefore Eq. (4) provides a robust and fast method to obtain a first estimate of the emitter location.

The solution proposed here, unlike the one described in [18], is based on a system that relies on one emitter and multiple receivers, with the algorithm being shaped on this assumption. The use of a single central emission pulse per cycle avoids the need for time scheduling between different transmitters. On the other hand, the system described in [18] requires a 65ms time allocation per emitter, reducing overall scalability. Moreover, the solution proposed here removes the requirement for the ultrasonic transmitters to emit signals with known periods and order, thus further reducing system complexity.

3.2. Localisation Improvement Via Optimization

There are several sources of noise and uncertainty in the system, therefore an optimisation procedure is used to minimise the effect on such factors on localisation accuracy. The algorithm described in Section 3.1 provides a good first estimate of the emitter location, but its accuracy is limited by the presence of noise and uncertainty limits. Therefore, the estimate obtained in Eq. (4) is used as first guess in an iterative optimization procedure aimed at improving localisation accuracy by adjusting the estimate $\boldsymbol{p}$. Such optimization problem can be mathematically expressed as

$$\boldsymbol{p_{estimated}} = argmin_{\widetilde{\boldsymbol{p}}_0}[S(\boldsymbol{p})] \tag{5}$$

$$S(\boldsymbol{p}) = \sum_{i=1}^{N}[(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2 - (t_0+t_{i0})^2 v_{sound}^2]^2 \tag{6}$$

A simple steepest descent algorithm is proposed to solve such problem. Note that Gauss-Newton or Levenberg-Marquardt are often preferred thanks to their superior convergence properties [21], but they require significantly more computational and memory resources. Moreover, simulation results indicate that such more advanced methods do not offer significant improvement compared to the simpler steepest descent method for the localisation scenario considered here.

The pseudo-code of the implemented steepest descent algorithm proposed for the localisation platform is shown in Table 2. Both the step size and the stopping criterion for these experiments were chosen heuristically. The step size was optimized to ensure convergence without requiring too many steps and the stopping criterion ensures that as soon as convergence occurs, the iterative minimiser stops. The constraint $k < k_{max}$ is included to terminate the iterative process so that there exists an upper bound for its execution time. In fact, very often (e.g. for real mobile robotic control) it is more important to get position estimates with a good rate, even at the expenses of accuracy. Moreover, it is worth noting that most of the benefits of steepest descent are realised in the first few iterations [23], therefore even a small value of k_{max} is sufficient to significantly improve accuracy with respect to the estimate provided by Eq. (4).

Table 2. Pseudo-code for solving problem Eqs. (5)-(6).

1. Initialise $\boldsymbol{p_{estimated}}(\boldsymbol{1})$ with results from Eq. (4) 2. Calculate step direction $\boldsymbol{step}(\boldsymbol{k}) = -\boldsymbol{\nabla S}(\boldsymbol{p_{estimated}}(\boldsymbol{k}))$ 3. Perform an optimisation step, with a predetermined step size $\boldsymbol{\alpha}$ $\boldsymbol{p_{estimated}}(\boldsymbol{k}+\boldsymbol{1}) = \boldsymbol{\alpha} \cdot \boldsymbol{step}(\boldsymbol{k}) + \boldsymbol{p_{estimated}}(\boldsymbol{k})$ 4. Check the stopping criterion for convergence 5. If the stopping criterion is not satisfied and $k < k_{max}$ repeat steps 2-5, otherwise return $\boldsymbol{p_{estimated}}(\boldsymbol{k})$

For the results shown in this paper we used $k_{max} = 40$ and the stopping criterion $|\boldsymbol{p_{estimated}}(k) - \boldsymbol{p_{estimated}}(k-1)| < 0.0002$. On the standard Arduino Mega 2560 platforms used for validation it takes on average 0.7 s to get the final estimate for emitter location, of which about 0.3 s are spent for running the minimisation algorithm.

4. Experimental Setup for Validation

To validate the performance of the proposed localisation system have been validated by implementing it and comparing its accuracy against a state-of-the-art motion tracking system, as described in this section.

4.1. Validation Setup

The processor unit utilised is an Arduino Mega 2560, which detects signals from the receivers and runs algorithm 1 and 2 for localisation. The output of this Arduino is connected to a Raspberry Pi to transmit all coordinates to a laptop to make a comparison with the coordinates provided by a motion tracking system composed of 8 Optitrack Prime 17W cameras. The emitter unit also utilizes an Arduino Mega 2560 board to trigger emission. Both units are connected via an RF transmitter/receiver (ERA-ARDUINO-S900) to trigger a new emission only once the localisation algorithm is finished. Motion capture results are captured at 100 Hz, whereas algorithm 2 takes about 0.7 s to run on the Arduino Mega 2560 board. To synchronise these two datasets, motion capture results are acquired and stored until the localiser has calculated a position, and then the nearest pair of coordinates from the 10 most recently acquired points is chosen to perform the comparison. Note that the robot used for moving the emitter has a speed of less than 1 cm/s, therefore such synchronisation strategy is accurate.

4.2. Description of the Hardware Implementation

The 12 receivers used are composed of US ceramic transducers (MCUSD16A40S12RO), which resonate at 40 kHz when detecting the pulse generated by the emitter. The received signal is then amplified 8-fold by a three-stage analogue amplifier and then converted to a square wave via a Schmidt-comparator (LM386). This signal processing electronics was inspired by the one described in [22]. The threshold value of the Schmidt-comparator was chosen to maximise range while avoiding the possibility of signal noise being amplified. For the setup described in this paper values between ~0.04 V and ~0.9 V provided a good trade-off between range and noise removal. This output of the Schmidt-comparator is then rectified and filtered by a passive Low Pass Filter at 15.9 kHz. Finally, a comparator (LM339-N) with cut off voltage is 1.9 V is used to generate the trigger signal to be transmitted to the processing unit.

The emitter is a composed by 13 US transducers (MCUSD16A40S12RO), with their outer metallic case removed to reduce the directionality properties. The support for such transducer was designed and 3D-printed so that an (almost) omnidirectional emission was achieved, as shown in Fig. 1. By doing this, the emitter can be considered a point source as required by the localisation algorithm described in Section 3. The 13 transducers are simultaneously pulsed using a microcontroller (PIC12F1822) generating a 40 kHz square wave, which is then amplified via a MOSFET amplifier and fed to the US emitters.

Experimental tests highlighted that the number of pulses affects the performance of the localisation system both in in terms of coverage and accuracy. More pulses result in increased coverage but decreased accuracy, and vice versa. As reported in Fig. 2, the mean overall error across the arena is proportional to the number of pulses. When a single pulse is emitted, the initial time t_0 will be identical for any active receiver. However, if there are two pulses and some receivers pick up the first pulse and some receivers pick up the second pulse, then t_{i0} may be different for these two sets of receivers, thus introducing additional uncertainty and increasing the localisation error. This effect may become significant for larger numbers of pulses. In fact, at 40 kHz, the time between pulses is 12.5 µs, which corresponds to 4.3 mm/pulse with a speed of sound $v_{sound} = 343\ m/s$. However, this uncertainty is significantly mitigated by the optimiser. On the other hand, the coverage of the system relies solely on the number of receivers picking up a signal at any given emission. Given the receiver circuits used, a weak reception may not always be immediately detected from the first pulse and could take multiple pulses from the same signal before detection is triggered. Therefore, the more pules being transmitted, the more likely a receiver will register a reception and the higher the coverage is, as shown in Fig. 3. The optimal number of pulses for the setup used for validation was determined to be five to ensure almost complete coverage while retaining good accuracy.

Fig. 1. 'Omnidirectional' US transducer array for the emitter.

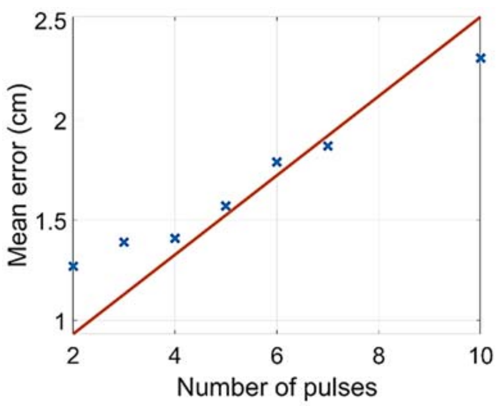


Fig. 2. Mean error (cm) vs. number of pulses.

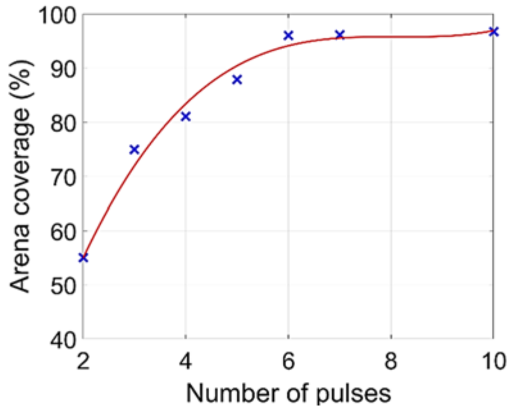


Fig. 3. Arena coverage (%) vs. number of pulses.

5. Results

Testing was done in a motion capture arena of approximately 4.3 m by 3.1 m, the emitter unit was placed on board a Create 2 Programmable Robot (iRobot, USA), that performed random movement around the arena for up to 1 hour per test. The results from both the localiser and motion capture were stored offline during this time. The system was tested at four different emitter heights from the floor: 280 mm, 355 mm, 457 mm and 592 mm, to prove its robustness. A summary of the results obtained of these experiments is reported in Table 3.

The 'pre-minimiser results' are obtained using Eq. (4), whereas results from the optimization routine described in Section 3 are referred as 'post-minimiser results'. Finally, given that the robot cannot move faster than approximately 1 cm/s and that algorithm 2 provide results every 0.7 seconds, any localisation estimate that is more than 15 cm apart from the last estimate can be considered as outlier and removed. Such outlier removal provides the set of 'post-filter results'. As can be seen in Table 3, localisation performance has similar trends across all the tested heights. In the following, only the results attained at 355 mm altitude are reported as illustrative examples.

Table 3. Localisation results at different heights.

Height (mm)	Mean Error *Post-Filt* *(cm)*	Points with error <1 cm *Post-Filt* (%)	Points with error < 3cm *Post-Filt* (%)	Coverage (%)
280	1.61	40.1	87.9	87.87
355	1.57	43.05	88.69	93.40
457	1.81	36.43	85.45	86.27
592	1.62	41.06	89.08	79.40

5.1. Pre-Minimiser

The pre-minimiser results obtained by Eq. (4) and shown in Fig. 4 are quite inaccurate. In fact, only 14.2 % of the results are within 1 cm of the true position, 47.8 % are within 3 cm and 81.5 % are within 10 cm. The mean error across all attained values is 11.2 cm. Such relatively high mean localisation error is significantly affected by the presence of extremely large errors (>30 cm) due to emitter reflections being picked up on the receiver units, a phenomenon that will be further discussed later.

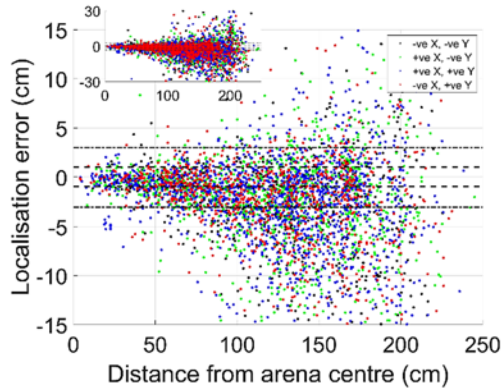


Fig. 4. Pre-Minimiser localisation error (cm) vs distance from arena centre, where different colours are used for the four quadrants of the arena. Inset plot shows a zoomed-out view of the localisation error (range ±30 cm).

Another noticeable trend is the monotonic increase of error spread as a function of distance from the centre of the arena. This is likely due to the number of receivers being involved in a positional calculation; as the robot moves away from the centre, some receivers go out of range and therefore less information is provided to the localisation algorithm.

Note that there exist spurious error clusters in certain quadrants, as highlighted in Fig. 5. Upon further inspection, these outliers were the results of some receivers picking up a reflected emission rather than the actual emission. In certain areas, this will happen in such a way that both algorithms will converge on an incorrect location. This could be mitigated by increasing delays between emission, logic filtering or reducing emission strength to decrease reflection likelihood.

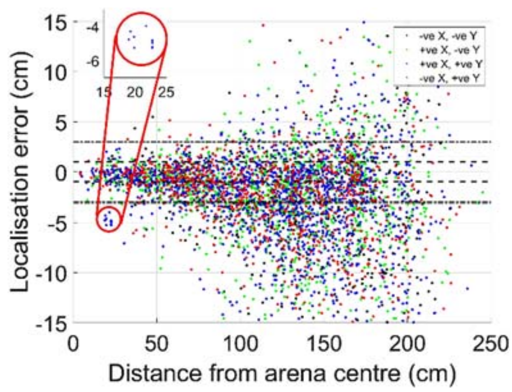


Fig. 5. Highlighted area of reflected reception.

To better quantify accuracy, the heat map shown in Fig. 6 was produced by discretizing the arena on a grid whose cells are 40 cm wide and reporting the median localisation error value for each cell. An 'Inf' value within a cell represents an area that has not been explored by the mobile robot during tests, and it is not taken into account when calculating means. The colour scaling is based on how accurate the data is.

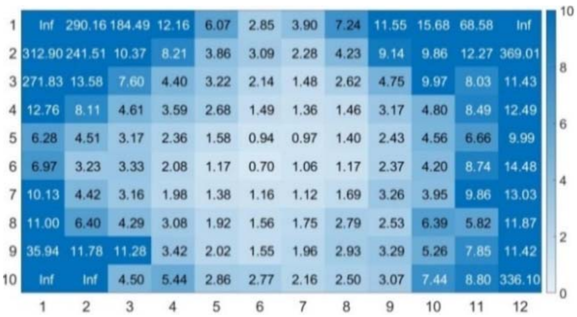


Fig. 6. Pre-Minimiser heatmap of the median localisation error (in cm) across the arena. Each cell represents a 40 cm×40 cm region in the arena,

5.2. Post Minimiser

After processing the results through the algorithm reported in Table 2, 41.7 % of the results are within 1cm of the true position, almost 3 times as many as the pre-minimised results. Moreover, 85.9 % of results are now within 3 cm and 96.4 % are within 10 cm of true position. Such improved performance translates to an average mean error of 3.33 cm. As shown in Fig. 7, the post minimiser results are more accurate throughout the arena, and the error does not significantly increase as the emitter moves away from the centre of the arena.

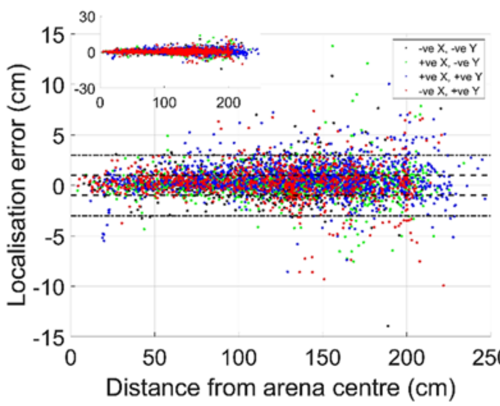


Fig. 7. Post-Minimiser localisation error (cm) vs distance from arena centre where different colours are used for the four quadrants of the arena. Inset plot shows a zoomed-out view of the localisation error (range ±30 cm).

The heat map shown in Fig. 8 demonstrates that the vast majority of results lie well within acceptable tolerances, with only 5 out of 120 cells having median errors above 5 cm.

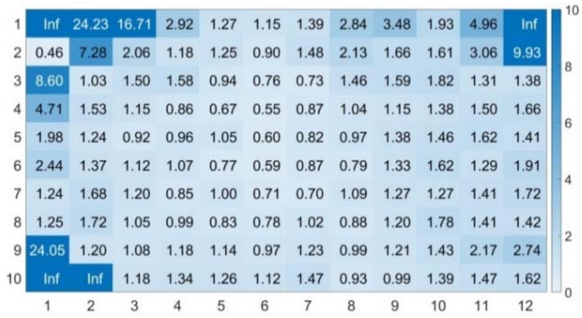


Fig. 8. Post-Minimiser heatmap of the arena, showing median localisation error (in cm) for 40 cm×40 cm cells.

5.3. Post Filter

When applying the final filter to remove any value that differs more than 15 cm from the previous estimate, the accuracy improves even further: 43.1 % of the results are within 1 cm of the true position, 88.7 % are within 3 cm and 99.5 % are within 10 cm. The overall average mean error drops to 1.57 cm as well. Fig. 9 reports the heat map related to these results.

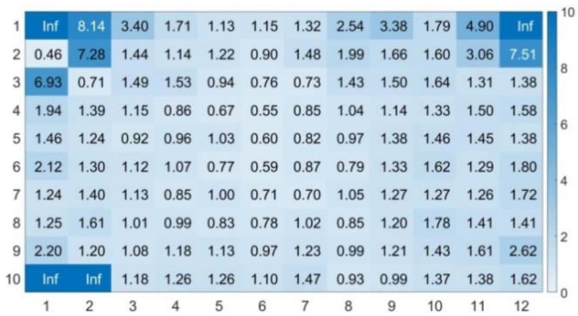


Fig. 9. Post-Filter heatmap of the arena, showing median localisation error (in cm) for 40 cm×40 cm cells.

Fig. 10 reports the median error (blue dots and lines) and the mean error (red dots and line) as functions of distance from the arena centre.

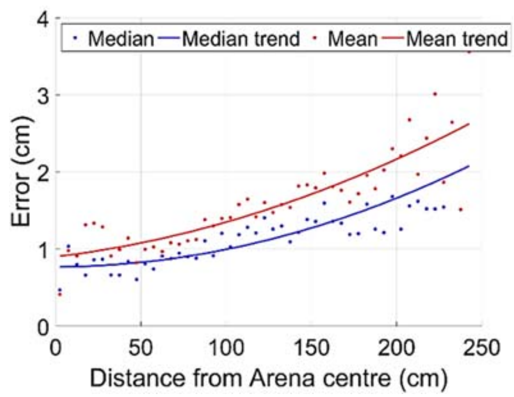


Fig. 10. Median error (red) and mean error (blue) vs distance from arena centre.

Both trends can be fitted by quadratic polynomials with the mean error showing some anomalies at 20 cm from arena centre, which result from the clustered reflections described earlier. The median result is far more robust to such reflection areas and shows a consistently lower error throughout the arena, indicating that the error distribution is skewed. The standard deviation of the results also increases with distance from arena centre as shown in Fig. 11, with an average standard deviation throughout the arena of 1.39 cm.

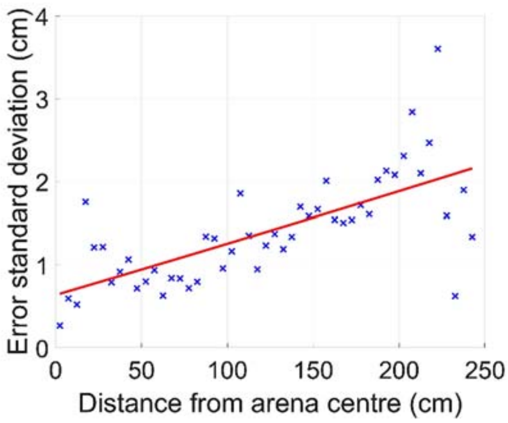


Fig. 11. Standard deviation of localisation errors vs distance from arena centre.

5.4. Localisation Summary

Table 4 summarises the results of each stage of processing, highlighting that the proposed minimisation and filtering algorithms significantly improve localisation performance.

Table 4. Summary of results obtained at various processing stages. E=Error, Min=Minimiser, Filt=Filter.

	Mean E	E<1 cm	E<3 cm	E<10 cm
Pre-Min	11.2 cm	14.2 %	47.8 %	81.5 %
Pre-Filt	3.33 cm	41.7 %	85.9 %	96.4 %
Post	1.57 cm	43.1 %	88.7 %	99.5 %

As previously mentioned, reflections played a major role in certain areas and significantly affected mean accuracy in these areas. One way in which this could be mitigated would be the addition of a constant time delay to ensure the dissipation of any remnant signals from previous pulses. Another, much more challenging solution would be the substitution of the Ultrasonic platform to UWB, removing any possible reflections. Usage in open outdoor areas would also remove any reflective areas. Another aspect to consider is the maximum scalability of the system. The current ultrasonic emitters have been tested to ranges of up to 12 m; however, the entire localisation systems accuracy has not been validated beyond the current setup.

5.5. Mobile Ground Robot Integration

To showcase the effectiveness of the proposed localisation platform, a path following demo for ground robots has been developed. To this end, the localisation unit was attached to a tracked robotic platform, as shown in Fig. 12. The integrated mobile robot was then given a series of waypoints that it was required to obtain in sequence to travel along a desired path.

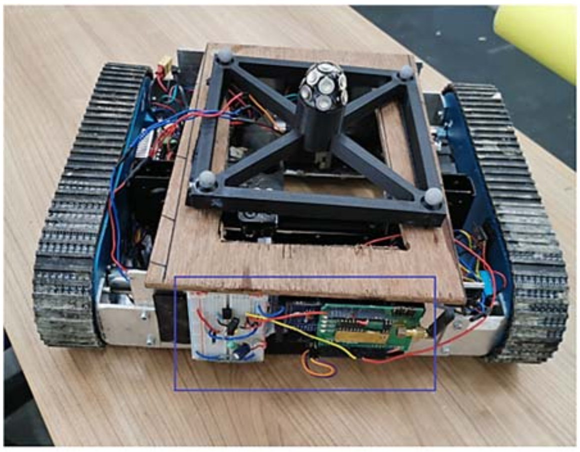


Fig. 12. Image of tracked robot integrated with localisation system. The blue box highlights the localisation unit, whereas the omnidirectional emitter surrounded by motion tracking markers was fitted to the top.

The proposed localisation system worked as an independent module that continually acquired positional information at the maximum localisation rate. The master controller of the robot would routinely request the most current (X, Y) coordinates from the localisation unit, which it would then utilise to determine the course of action. Orientation information was acquired using an inertial measurement unit board (Arduino A000070). Provided the most recent position and orientation, the system corrected orientation on the spot to intersect the rover path with the next way point before initiating translational movement in a straight line. Whilst moving towards this waypoint, the system continually performed regular heading checks based on the position and orientation data being acquired periodically. By using this method, the system would reach the waypoint and then proceed to the next waypoint using the same procedure.

It was unrealistic to assume exact heading and waypoint acquisition, therefore tolerances were imposed. Tolerances in heading were variable depending on distance to waypoint, ranging from 3° when the system was over 500 mm from the waypoint, to 15° when the system was less than 70 mm from the waypoint. Waypoint tolerance was set to a constant 25 mm. These values were tuned to ensure the optimum balance between accurate path following and smoothness of movement.

For the purpose of the demo, the system was tasked with following a set of waypoints representing a square whose sides were 170 cm long. Snapshots of the rover achieving the corners of this planned path are shown in Fig. 13.

The rover was also fitted with an array of motion capture markers, as shown in Fig. 12, which enabled more accurate rover path acquisition, using the same technique described in Section 4.1. The results from this test are shown in Fig. 14. The plot consists of over 33,000 motion capture data points acquired over the course of 10 minutes of path following. The red circles shown have a radius of 25 mm and represent the waypoints with the associated tolerance. As can be seen from the results, the waypoints were successfully reached within the specified tolerance. The successful completion of this demo demonstrates that the low-cost ultrasound-based localisation system proposed in this paper is suitable for robotic applications such as path following.

Fig. 13. Aerial snapshots of the mobile robot following the predetermined waypoints.

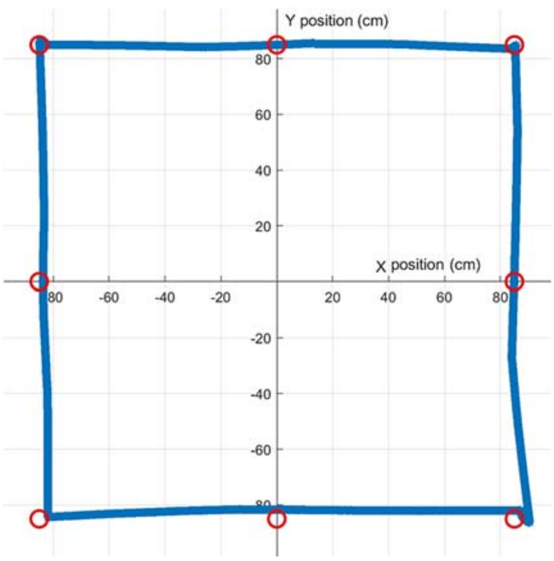


Fig. 14. Plot of the mobile unit's exact position shown by the blue line, where the waypoints are highlighted as the centre point with a circle of radius equal to goal tolerance.

6. Conclusions

In this paper an inexpensive yet accurate ultrasound-based localisation system is proposed. The total cost of the components is around £100, the biggest share of which are two Arduino Mega 2560 boards used for trigger emission and for processing data at the receivers end. The system allows localisation of 89 % of the 4.3 m × 3.1 m arena with an accuracy of less than 3 cm and 43 % with an accuracy of less than 1 cm. The system has been proven to be scalable between 280 mm and 592 mm of height, without requiring any change in the experimental setup. The proposed localisation system is generic and can be used for a variety of applications [24]. For example, it was integrated with a ground robot to show that it can be successfully used to enable the robot to perform waypoint-based path following.

Given the range of the ultrasound transceivers used, the system should theoretically perform well in arenas up to 12×12×12 m in size. Optimisation of receiver location and orientation and use of more powerful transducers would also allow better performance in larger arenas if needed. Finally, more advanced filtering approaches may be developed to improve robustness with respect to spurious reflections.

Acknowledgements

This research was supported by Apadana Management 3 Ltd. The authors also wish to thank colleagues from the University of Liverpool who provided insight and expertise that greatly assisted the research.

References

[1]. H. Cheng, H. Chen, Y. Liu, Topological Indoor Localization and Navigation for Autonomous Mobile Robot, *IEEE Transactions on Automation Science and Engineering*, Vol. 12, Issue 2, 2014, pp. 729-738.

[2]. C. C. Ward, K. Iagnemma, A Dynamic-Model-Based Wheel Slip Detector, *IEEE Transactions on Robotics*, Vol. 24, Issue 4, 2008, pp. 821-832.

[3]. L. Paya, A. Gil, O. Reinoso, A State-of-the-Art Review on Mapping and Localization of Mobile Robots Using Omnidirectional Vision Sensors, *Journal of Sensors*, Vol. 2017, 2017, pp. 1-20.

[4]. A. A. Khudhair, S. Q. Jabbar, M. Q. Sulttan, D. Wang, Wireless Indoor Localization Systems and Techniques: Survey and Comparative Study, *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 3, Issue 2, 2016, pp. 392-409.

[5]. AT&T Laboratories Cambridge, The Bat Ultrasonic Location System, AT&T, 2002. [Online]. Available: https://www.cl.cam.ac.uk/research/dtg/attarchive/bat/ . [Accessed 23 August 2018].

[6]. L. Batistic, M. Tomic, Overview of Indoor Positioning System Technologies, *International Convention on Information and Communication Technology,*

Electronics and Microelectronics (MIPRO), Issue 41, 2018, pp. 473-478.
[7]. Marvelmind Robotics, Starter Set – HW v4.9, August 2018. [Online]. https://marvelmind.com/product/starter-set-hw-v4-9-plastic-housing/. [Accessed August 2018].
[8]. Seeed, UWB PK-1000 Standard Edition, Pudutech, 2017. [Online]. https://www.seeedstudio.com/UWB-PK-1000-Standard-Edition-p-2918.html. [Accessed July 2018].
[9]. H. Liu, H. Darabi, P. Banerjee, J. Liu, Survey of Wireless Indoor Positioning Techniques and Systems, *IEEE Transactions on Systems, Man, and Cybernetics,* Vol. 37, Issue 6, 2007, pp. 1067-1080.
[10]. S. Tekinay, E. Chao, R. Richton, Performance benchmarking for wireless location systems, *IEEE Communications Magazine*, Vol. 36, Issue 4, 1998, pp. 72-76.
[11]. E. Tinti, Metodi e Stumenti per il Posizionamento Indoor, Università di Bologna-Campus Di Cesena - Scuola Di Scienze, Bologna, *Anno Accademico*, 2014.
[12]. S. Ingram, D. Harmer, M. Quinlan, Ultrawideband indoor position systems and their use in emergencies, in *Proceedings of the Position Location and Navigation Symposium*, Monterey, CA, USA, 2004.
[13]. P. Pannuto, B. Kempke, L.-X. Chuo, D. Blaauw, P. Dutta, Harmonium: Ultra Wideband Pulse Generation with Bandstitched Recovery for Fast, Accurate, and Robust Indoor Localization, *ACM Transactions on Sensor Networks*, Vol. 14, Issue 2, 2018, p. Article 11.
[14]. R. F. Brena, P. Garcia-Vazquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, J. Fangmeyer Jr., Evolution of Indoor Positioning Technologies: A Survey, *Journal of Sensors*, Vol. 2017, 2017, Article ID 2630413.
[15]. A. Yassin, *et al*., Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications, *IEEE Communications Surveys & Tutorials,* Vol. 19, Issue 2, 2017, pp. 1327-1246.
[16]. G. Mao, B. Fidan, Localization Algorithms and Strategies for Wireless Sensor Networks, *Information science Reference-Imprint of IGI Publishing,* 2009.
[17]. R. Kaune, J. Horst, W. Koch, Accuracy Analysis for TDOA Localization in Sensor Networks, in *Proceedings of the 14^{th} International Conference on Information Fusion*, Chicago, Illinois, USA, 2011.
[18]. U. Yayan, H. Yucel, A. Yazici, A Low Cost Ultrasonic Based Positioning System for the Indoor Navigation of Mobile Robots, *Journal of Intelligent & Robotic Systems*, Vol. 78, Issue 3-4, 2015, pp. 541-552.
[19]. Y. Zhou, An efficient least-squares trilateration algorithm for mobile robot localization, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 3474-3479.
[20]. K. W. Cheung, H. C. So, W.-K. Ma, Y. T. Chan, A Constrained Least Squares Approach to Mobile Positioning, *EURASIP Journal on Applied Signal Processing*, 2006, pp. 1-23.
[21]. S. J. Wright, J. Nocedal, Numerical Optimization, *Springer*, 1999.
[22]. E. Electronics, Engineeringshock Electronics, 2018. [Online]. http://www.engineeringshock.com/ultrasonics.html. [Accessed 23 May 2019]
[23]. R. M. Freund, The Steepest Descent Algorithm for Unconstrained Optimization and a Bisection Line-search Method, *Massachusetts Institute of Technology,* Massachusetts, 2004.
[24]. A. J. Burns, S. Fichera, P. Paoletti, A low cost ultrasound-based localisation system, in *Proceedings of the 1^{st} IFSA Frequency & Time Conference (IFTC'19)*, Tenerife (Canary Islands), 25-27 September 2019, pp. 21-27.

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 31-35

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Research on the Blind Zone Positioning in Dual-satellite Passive Location System

[1, 3, *] Yanrong XUE, [1, 2] Yuwei LI, [1, 2] Ruiqiong CHEN, [1, 2] Duosheng FAN and [1, 2, 3] Xiaohui Li

[1] National Time Service Center, Chinese Academy of Sciences, No. 3 East Shuyuan Road, Lintong, Xi'an, Shaanxi, 710600, China

[2] Key Lab. of Time & Frequency Primary Standards, Chinese Academy of Sciences, No. 3 East Shuyuan Road, Lintong, Xi'an, Shaanxi, 710600, China

[3] Key Lab. of Precision Navigation and Timing Technology, Chinese Academy of Sciences, No. 3 East Shuyuan Road, Lintong, Xi'an, Shaanxi, 710600, China

[1] Tel.: 029-83890851, fax: 029-83890326
E-mail: xueyanrong@ntsc.ac.cn

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: Dual-satellite passive location system is a passive location method based on time difference of arrival and frequency difference of arrival. In order to improve the positioning accuracy, the position differential and pseudo-range differential method of satellite navigation system is used to correct the positioning error of this system, and the positioning accuracy after the correction is about 5 km. But there is one problem when we use the method. That is, basing on the result of positioning error distribution, positioning error in the band about ±100 km in the direction of the line between the two satellites using time difference of arrival and frequency difference of arrival positioning is very big, and positioning accuracy is very poor or the emitter can't be positioned in that area, which is called blind zone. In order to solve the difficult positioning problem in the blind zone, the reason for the positioning blind zone is given and a new method of multi-epoch time difference positioning is proposed. The results of solving method and simulation show that high positioning accuracy can be obtained by using the multi-epoch time difference method for the positioned emitter in the blind zone, which can effectively solve the problem of blind zone in positioning area.

Keywords: Time difference of arrival, Frequency difference of arrival, Multi-epoch, Time difference, Blind zone.

1. Introduction

Dual-satellite passive location system is based on joint positioning technology of time difference of arrival (TDOA) and frequency difference of arrival (FDOA), simultaneously the time difference and frequency difference information of the same emitter on the ground getting to the two satellites is measured to obtain the location of the signal source. The principle of dual-satellite passive location system basing on TDOA and FDOA is as shown in Fig. 1, where the emitter is T, the two satellites are S1 and S2, the receiving station is A.

Initially, time difference of arrival and direction measurement were used to achieve the positioning [1-2], then the time difference and the frequency difference of positioning technology was researched [3]. Domestic scholars have also carried out algorithm research of the time difference and the frequency difference dual-satellite passive location [4-7]. The positioning accuracy characteristics of the dual-satellite passive location system were also investigated, and positioning accuracy curve is obtained [8]. The influence of relative positions of dual satellites, measurement error of TDOA and FDOA, measurement error of satellite position and velocity on the geolocation accuracy of high and low orbit dual satellites is demonstrated [9].

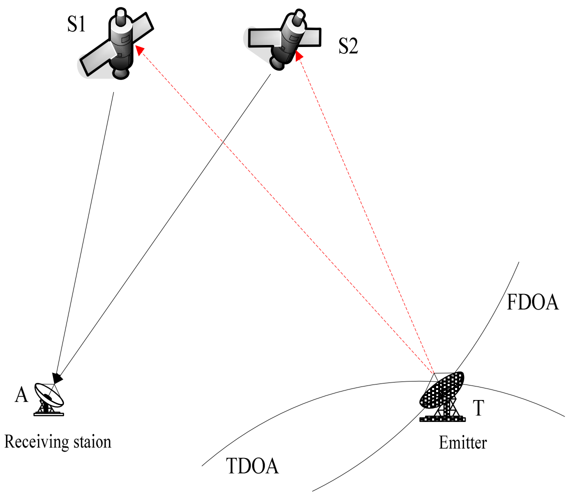


Fig. 1. The principle of dual-satellite passive location system basing on TDOA and FDOA.

Nowadays, positioning accuracy for dual-satellite passive location system can achieve 10 km or so by using current algorithm, which is relatively poor. Therefore differential method is applied to three satellites passive location by some scholars [10], but almost no scholars study differential techniques applying to dual-satellite passive location. The position differential technology of GPS positioning error correction method applied to positioning error correction technology of the dual-satellite passive location is researched for the first time [11]. But the accuracy is poor, and we can use the method unless the distance between the emitter and reference station is within hundreds of kilometer and the GDOP value of the emitter and reference station is close. So combining with pseudo-range differential in the application of GPS, the pseudo-range differential method in Search and Rescue with two satellites is also proposed innovatively [12], and its positioning accuracy is analysed, which has taken a critical step for the study of the passive location and also improves the passive positioning accuracy greatly, but the accuracy is not high enough, and there are some limitations in the application of this method, particularly, which can't be used for some area.

Moreover, basing on the result of positioning error distribution in dual-satellite passive location system, shown in Fig. 2(a) and Fig. 2(b), positioning error in the band about ±100 km in the direction of the line between the two satellites by using TDOA and FDOA positioning equation is so big, and positioning accuracy is very poor or emitter can't be located in that area, which is called blind zone, particularly at the inner sub-satellite point area. In order to ensure that there is no blind zone in the dual-satellite passive location system, it is necessary to solve the problem of positioning for blind zone effectively.

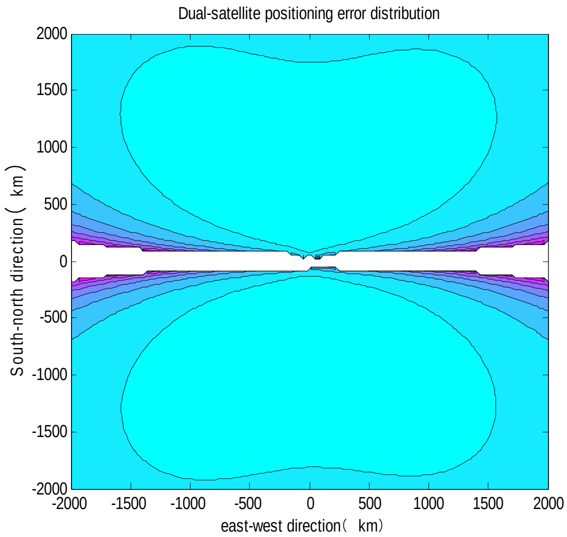


Fig. 2 (a). The result of positioning error distribution in dual-satellite passive location system.

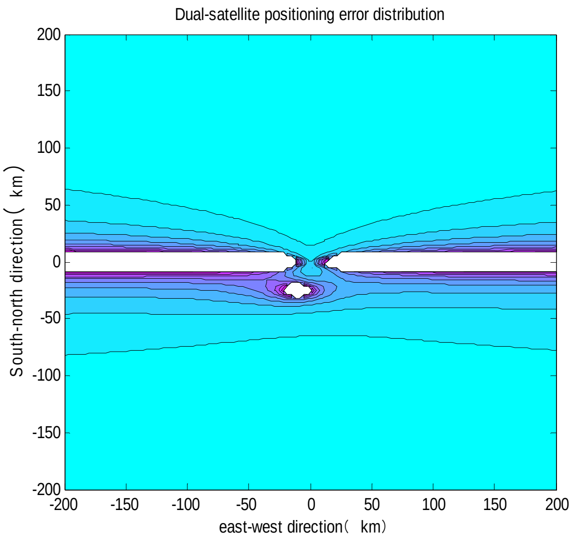


Fig. 2 (b). The result of positioning error distribution in dual-satellite passive location system.

Therefore, we found that the frequency difference error is very big, and it is not proper to use a separate frequency difference location system. While in the inner sub-satellite point area the positioning error caused by the time difference error is the smallest one and the time difference error in the blind zone is smaller than that of the non-blind-zone. A possibility is provided by these analyzes to solve the blind zone location problem by using the TDOA system.

2. The Principle of Multi-epoch Time Difference Positioning

As the satellite motion is continuous, a number of time difference measurements value can be obtained if measuring time difference of arrival of the measured emitter for several times continuously through the passive signal of dual-satellite. Orbit position of the satellites for several epochs can be determined at the same time by using satellite navigation receivers. Based on the principle of time difference measurement, the time difference equation for several epochs can be obtained, then position solution can be obtained, which is called the multi-epoch time difference method.

A set of three positioning equations about time difference can be obtained, assuming three successive epochs,

$$c\Delta t_1 = \sqrt{(x_1^{s1}-x)^2+(y_1^{s1}-y)^2+(z_1^{s1}-z)^2} - +\sqrt{(x_1^{s2}-x)^2+(y_1^{s2}-y)^2+(z_1^{s2}-z)^2} \quad (1)$$

$$c\Delta t_2 = \sqrt{(x_2^{s1}-x)^2+(y_2^{s1}-y)^2+(z_2^{s1}-z)^2} - \sqrt{(x_2^{s2}-x)^2+(y_2^{s2}-y)^2+(z_2^{s2}-z)^2} \quad (2)$$

$$c\Delta t_3 = \sqrt{(x_3^{s1}-x)^2+(y_3^{s1}-y)^2+(z_3^{s1}-z)^2} - \sqrt{(x_3^{s2}-x)^2+(y_3^{s2}-y)^2+(z_3^{s2}-z)^2} \quad (3)$$

In the above formula, superscript S1 and S2 represent the two satellites, the subscript 1, 2 and 3 represent the epoch.

For the above three formulas, partial derivatives on *x*, *y*, *z* respectively, 9 partial derivatives can be obtained, composing of the Jacobian matrix F_x, denote as respectively

$$F_X = \begin{bmatrix} F_{1x} & F_{1y} & F_{1z} \\ F_{2x} & F_{2y} & F_{2z} \\ F_{3x} & F_{3y} & F_{3z} \end{bmatrix}$$

Then,

$$F_{1x} = \frac{x - x_1^{s1}}{r_{11}} - \frac{x - x_1^{s2}}{r_{21}} \quad (4)$$

$$F_{1y} = \frac{y - y_1^{s1}}{r_{11}} - \frac{y - y_1^{s2}}{r_{21}} \quad (5)$$

$$F_{1z} = \frac{z - z_1^{s1}}{r_{11}} - \frac{z - z_2^{s2}}{r_{21}} \quad (6)$$

$$F_{2x} = \frac{x - x_2^{s1}}{r_{12}} - \frac{x - x_2^{s2}}{r_{22}} \quad (7)$$

$$F_{2y} = \frac{y - y_2^{s1}}{r_{12}} - \frac{y - y_2^{s2}}{r_{22}} \quad (8)$$

$$F_{2z} = \frac{z - z_2^{s1}}{r_{12}} - \frac{z - z_2^{s2}}{r_{22}} \quad (9)$$

$$F_{3x} = \frac{x - x_3^{s1}}{r_{13}} - \frac{x - x_3^{s2}}{r_{23}} \quad (10)$$

$$F_{3y} = \frac{y - y_3^{s1}}{r_{13}} - \frac{y - y_3^{s2}}{r_{23}} \quad (11)$$

$$F_{3z} = \frac{z - z_3^{s1}}{r_{13}} - \frac{z - z_3^{s2}}{r_{23}} \quad (12)$$

In the above formula, subscript *k* of r_{ki} denotes the corresponding satellite *k* = 1, 2, and the subscript *i* denotes the corresponding epoch. Its composition matrix is

$$F_x = \begin{bmatrix} F_{1x} & F_{1y} & F_{1z} \\ F_{2x} & F_{2y} & F_{2z} \\ F_{3x} & F_{3y} & F_{3z} \end{bmatrix}$$

Assuming the time difference measured is σ_1, σ_2, σ_3. According to this multi-epoch time difference method, the position of the emitter can be obtained by using the classical Newton iterative method in the Cartesian coordinate system.

3. Experiment Results

Effective satellite orbit data are used for this experiment, three simulation experiments results are given as following:

Simulation experiment 1: Let the satellite height be *H*, *H* is less than 1000 km, x_{11}=125, y_{11}=30.1. The distance between x_{i2} and x_{i1} is about 100 km, the distance between y_{i2} and y_{i1} is about 300 m, *i*=1, 2, 3. Every two epochs are more than 200 km, corresponding satellite orbit of two satellites and the resulting positioning error of emitter for three time epoch as follows:

Epoch *i*: sat1(i)=[x_{i1} y_{i1} H];

sat2(i)=[x_{i2} y_{i2} H]; $i = 1, 2, 3$.

The positioning error of the emitter after solution Δx_1 Δy_1 Δz_1 can be obtained.

Simulation experiment 2: Let the satellite height be *H*, *H* is less than 1000 km, x_{11}=125, y_{11}=30.1. The distance between x_{i2} and x_{i1} is about 200 m, the distance between y_{i2} and y_{i1} is about 300 m *i*=1, 2, 3. Every two epochs are less than 100 km, corresponding satellite orbit of two satellites and the resulting positioning error of emitter for three time epochs is similar with experiment simulation 2.

The positioning error of the emitter after solution:

$\Delta x_1 < \Delta x_2 < 1$ km
$\Delta y_1 < \Delta y_2 < 1$ km
$\Delta z_1 < \Delta z_2 < 1$ km

Simulation experiment 3: Let the satellite height be *H*, *H* is less than 1000 km, the satellite ephemeris error is 10 m (accuracy or orbit accuracy measured from the onboard navigation receiver), the two satellites orbit are the same with simulation experiment 2. The resulting positioning error of emitter for three time epoch as follows:

$\Delta x_1 \langle \Delta x_3 \langle \Delta x_2 \langle 1\text{km}$
$\Delta y_1 \langle \Delta y_3 \langle \Delta y_2 \langle 1\text{km}$
$\Delta z_3 \langle \Delta z_1 \langle \Delta z_2 \langle 1\text{km}$

From the above simulation experiments can be shown:

1) The multi-epoch the time difference method can solve the localization problem for the blind zone under the satellite, and the positioning accuracy is high.

2) The time space between epoch chosen should be longer appropriately, positioning effect will be better.

To describe the positioning error distribution, the above equations are transformed by using the geographic coordinate system. In the geographical coordinate system, the position of the emitter can be set to $(x, y, 0)$, the subscript (e, n, u) denotes the north-east-day coordinate system of the geographic coordinate system, and the coordinates of the satellite can be set as $(x_{ei}^k, y_{ni}^k, H), (k=1,2, i=1,2)$, k is the satellite number and i is the epoch number. In the geographical coordinate system, assuming that the satellite height is *H*, since there are only two unknowns, only two epochs are needed to determine the position of the emitter.

Two epochs are used by monitoring the emitter during satellite motion, let

$$r_i^k = \sqrt{(x - x_{ei}^k)^2 + (y - y_{ni}^k)^2 + H^2}$$
$$k=1,\ 2,\ i=1,\ 2$$

The positioning equation of multi-epoch time difference method is expressed as

$$c\Delta t_1 = \sqrt{(x - x_{e1}^2)^2 + (y - y_{n1}^2)^2 + H^2} - \sqrt{(x - x_{e1}^1)^2 + (y - y_{n1}^1)^2 + H^2}, \qquad (13)$$

$$c\Delta t_2 = \sqrt{(x - x_{e2}^2)^2 + (y - y_{n2}^2)^2 + H^2} - \sqrt{(x - x_{e2}^1)^2 + (y - y_{n2}^1)^2 + H^2}, \qquad (14)$$

For the above equation, the classical Newton iterative method is used to solve it here. The solving equation is as follows

$$F_1 = \sqrt{(x - x_{e1}^2)^2 + (y - y_{n1}^2)^2 + H^2} - \sqrt{(x - x_{e1}^1)^2 + (y - y_{n1}^1)^2 + H^2} - c\Delta t_1 \qquad (15)$$

$$F_2 = \sqrt{(x - x_{e2}^2)^2 + (y - y_{n2}^2)^2 + H^2} - \sqrt{(x - x_{e2}^1)^2 + (y - y_{n2}^1)^2 + H^2} - c\Delta t_2 \qquad (16)$$

For the above two expressions, the partial derivative on *x*, *y*. The subscript (e, n, u) represents the North-east-day coordinate system of the geographic coordinate system. Four partial derivatives can be derived from the procedure, composing of the Jacobian matrix F_x, denoting as respectively.

$$F_{1x} = \frac{x - x_{e1}^2}{r_1^2} - \frac{x - x_{e1}^1}{r_1^1} \qquad (17)$$

$$F_{1y} = \frac{y - y_{n1}^2}{r_1^2} - \frac{y - y_{n1}^1}{r_1^1} \qquad (18)$$

$$F_{2x} = \frac{x - x_{e2}^2}{r_2^2} - \frac{x - x_{e2}^1}{r_2^1} \qquad (19)$$

$$F_{2y} = \frac{y - y_{n2}^2}{r_2^2} - \frac{y - y_{n2}^1}{r_2^1} \qquad (20)$$

Matrix

$$F_x = \begin{bmatrix} F_{1x} & F_{1y} \\ F_{2x} & F_{2y} \end{bmatrix}$$

Assuming the time difference measured is σ_{t1}, σ_{t2}.

The position solution of the emitter can be obtained by the classical Newton Iteration method.

The simulation results for the blind zone are shown in Fig. 3(a) and Fig. 3(b).

From the above figure, high positioning accuracy can be obtained for the blind zone by using multi-epoch time difference method, which can solve the problem of positioning in blind zone. For practical applications, angle of arrival in the passive direction given by the signal can be used to determine roughly whether the emitter is located in the region which can't be located, which can also be determined by the threshold of the GDOP value. If it is the region which can't be located, using multi-epoch time difference location method, otherwise, using time difference frequency and time difference location method.

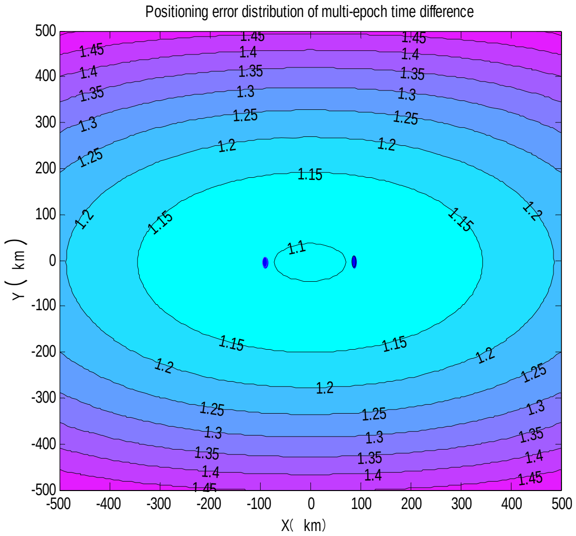


Fig. 3 (a). The positioning error distribution of blind zone by multi-epoch time difference method.

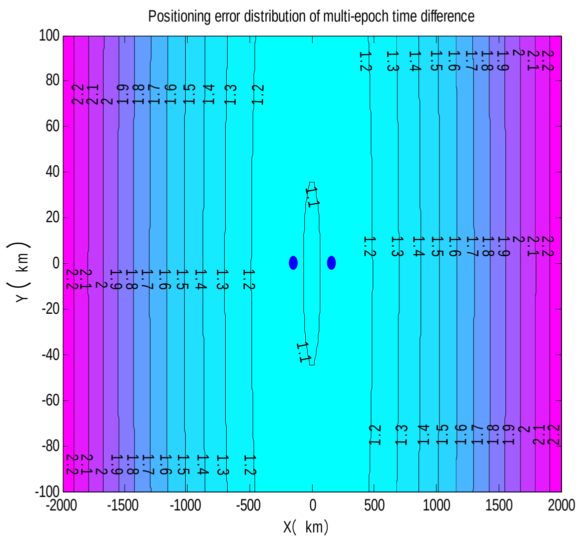


Fig. 3 (b). The positioning error distribution of blind zone by multi-epoch time difference method.

4. Conclusions

In this paper, the cause of blind zone generating is analyzed. In order to solve the problem of positioning in blind zone, a new method of multi-epoch time difference method is put forward to solve the problem of blind zone localization and make up for the defect of TDOA and FDOA location system. The results of resolution and simulation show that high positioning accuracy can be obtained for the blind zone by using multi-epoch time difference method, which can effectively solve the problem of blind zone in the location area.

Acknowledgements

Our thanks to the Support of Youth Innovation Promotion Association Chinese Academy of Sciences.

References

[1]. Wegner L. H., On the accuracy analysis of airborne techniques for passively locating electromagnetic emitters, Report R-722-PR, *Rand Corp., National Technical Information Service as ASTIA D.C. AD* 729 767, 1971.

[2]. Torrieri D. J., Statistical theory of passive location systems, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-20, Issue 2, 1984, pp. 183-198.

[3]. Ho K. C., Chan Y. T., Geolocation of a known altitude object from TDOA and FDOA measurements, *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 33, Issue 3, 1997, pp. 770-783.

[4]. Zhang Y., Sheng W. D., Guo F. C., *et al.*, Low orbit dual-satellites location algorithm and its precision analysis, *Journal of Chinese Inertial Technology*, Vol. 15, Issue 2, 2007, pp. 188-192.

[5]. Guo F. C., Fan Y., A method of Dual-Satellites geolocation using TDOA and FDOA and its precision analysis, *Journal of Astronautics*, Vol. 29, Issue 4, 2008, pp. 1381-1386.

[6]. Guo F. C., Fan Y., Analysis of dual-satellites geo-location method using TDOA and FDOA, *Aerospace Electron Confront*, Vol. 6, 2006, pp. 20-23.

[7]. Sun G. L., Shen S. T., Ding Z. M., *et al.*, An integration method for passive RDSS and DNS, *Acta Aeronautica et Astronautica Sinica*, Vol. 27, Issue 4, 2006, pp. 682-686.

[8]. Liu Yang, Liu Xiaobo, Analysis of Location Precision in Dual-satellite Passive Location System Based on GDOP, *Journal of Nanchang Hangkong University*, Vol. 31, Issue 4, 2017, pp. 7-13.

[9]. Di Hui, Pan Jinbo, Zhang Guoqiang, Zhao Qinguang, Geolocation Accuracy of High and Low Orbit Satellite by using TDOA and FDOA, *Aerospace Shanghai*, Vol. 35, Issue 8, 2018, pp. 54-59.

[10]. Liu J. Y., Kang G. H., Yue Y. Z., *et al.*, Integrated navigation system of three-star passive Beidou /SINS based on pseudo-range difference, *Journal of Chinese Inertial Technology*, Vol. 15, Issue 4, 2007, pp. 445-448.

[11]. Yanrong Xue, Xiaohui Li, Longxia Xu, *et al.*, Research on Position Differential Method of Dual-satellites TDOA and FDOA in Passive Location System, in *Proceedings of the IEEE International Frequency Control Symposium Proceedings*, Baltimore, USA, 21-24 May 2012, pp. 481-485.

[12]. Yanrong Xue, ShaoJun Feng, Washington Yotto Ochieng, *et al.*, The Improvement of the Positioning Accuracy in Search and Rescue with Two Satellites, in *Proceedings of the 7th China Satellite Navigation Conference (CSNC'16)*, Chang Sha, China, 18-20 May 2016, pp. 255-261.

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 36-41

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019

http://www.sensorsportal.com

Research on Dynamic Steering Method of Atomic Clock Based on Residual Correction

[1, 2, *] Yuwei Li, [1, 2] Wenli Wang, [1, 2] Yanrong Xue, [1, 2] Ruiqiong Chen and [1, 2] Ya Liu

[1] National Time Service Center, Chinese Academy of Sciences, China

[2] Key Lab. of Time & Frequency Primary Standards, Shaanxi, China

[1] Tel.: + (086)83890447, fax: (086)83890326

E-mail: liyuwei9999@163.com

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: Precision time-frequency signal generation and maintenance technology is an important foundation of navigation and positioning, which provides actual physical signal and precise time scale for navigation and positioning. In order to generate and maintain the frequency accuracy, frequency stability and synchronization accuracy of time-frequency signals, combined with the characteristics of atomic clock, a method of dynamic atomic clock control by residual correction is proposed. The method greatly improves the frequency accuracy and long-term frequency stability of signals without destroying the short-term stability. The result of synchronization is at least twice as much as that of traditional methods.

Keywords: Precision time-frequency signal, Dynamic atomic clock control, Residual correction, Signal maintenance technology, Minimum control strategy control technology.

1. Introduction

Precision time-frequency signal generation and maintenance technology is an important foundation of navigation and positioning, which provides actual physical signal and precise time scale for navigation and positioning. The performance of time-frequency signal directly affects the positioning accuracy and timing precision of navigation and communication systems.

Frequency deviation and frequency drift exist in the output signal of the free-running atomic clock, which results in the decrease of frequency accuracy, stability and synchronization accuracy. Therefore, it is necessary to control the atomic clock so as to stabilize the output signal of the atomic clock in a certain range and achieve high-precision synchronization.

This paper mainly studies the high precision time synchronization method of atomic clock. The precise time difference prediction method based on least square method, Kalman filter and the other control method of atomic clock are deeply analyzed. Combined with the characteristics of atomic clock, a method of dynamic atomic clock control by residual correction is proposed. The method uses the residual for correcting time difference model to improve the prediction accuracy. When the reference is available, the parameters, such as the attenuation factor and control period, are adjusted according to the time difference prediction in real time; when the reference is not available, an atomic clock autonomous maintenance algorithm is designed to keep the signal synchronized without the assistance of reference source from atomic clock or some external signal.

2. Main Steering Method

Usually, the atomic clock is used to generate high-precision time-frequency signals and system time. Due to the frequency deviation and frequency drift of the freely running atomic clock, the frequency accuracy of the output signal decreases. Therefore, the atomic clock needs to be controlled to stabilize the atomic clock output signal within a certain range.

The traditional atomic clock steering methods can be classified into two categories including phase adjustment and frequency adjustment.

The phase adjustment method is achieved by monitoring the time difference between the atomic clock and the reference in real time. When the time difference reaches the threshold value, we should adjust atomic clock so that phase can jump directly to zero. Then the atomic clock runs freely until the threshold is reached again, and the phase is still adjusted. The phase adjustment can change phase quickly but it can't change the frequency characteristics of the signal itself. The phase adjustment method will cause phase jump, destroy the frequency stability of the signal and affect the normal use of the signal.

The frequency adjustment method is based on fully understanding the performance of the atomic clock, using a fix-predetermined frequency offset or dynamic frequency offset to adjust the atomic clock. The method allows the atomic clock to run freely to the threshold, then we should adjust the frequency offset, and make it run in the opposite direction with the new clock speed until the other threshold is reached. In that time, adjustment must be cancelled.

At present, the most used steering method is combine the phase adjustment with frequency adjustment together. The method can keep phase jump continuously, and change tendency of frequency offset quickly and improve the performance of atomic clock.

3. Dynamic Steering Method Based on Residual Correction

The steering methods mentioned before are not run automatically and intelligently. Therefore, the steering method is proposed with running automatically and intelligently which is called dynamic steering method based on residual correction.

The key idea of steering method is to predict the time difference of the next state based on the change law of historical time difference data and determine the control parameters by the forecast value, such as control period, attenuation factor and others. Therefore, the steering method is closely related to the time difference forecast. The method contains two aspects which are residual error correction method and dynamic steering method.

3.1. Residual Error Correction Method

There are many factors that affect the accuracy of time difference prediction include frequency stability, frequency drift characteristics, time difference forecast duration and time difference data quality. Time difference forecasting methods commonly used include polynomial method, grey model estimation method, neural network, etc. Due to the error of alignment link and the data processing method, it is impossible to completely match any fitted model with the true value, and there is always a deviation called the residual error between the predicted value and the true value. The residual error is usually represented with $\hat{\varepsilon}$, and the residual error visually reflects whether the fitting model is suitable.

The residual error can't be zero because of the existence of random error. The aim of model fitting is reducing the residual value constantly. The average of the residuals represents the overall offset of the data during the current control period, and reflects the predict precise of time difference fitting model. If the time difference model was corrected with the mean value of the residual, the predict value can be made closer to the true value.

Residual error correction method is a new method that converts residual to frequency deviation, and corrects it into the time difference model, and establishes a new model for the time difference forecast to reduces the residual continuously. Therefore, the residual value can change in a convergence trend and the accuracy of the forecast can improve. The flow chart of residual error correction method is as shown in Fig. 1.

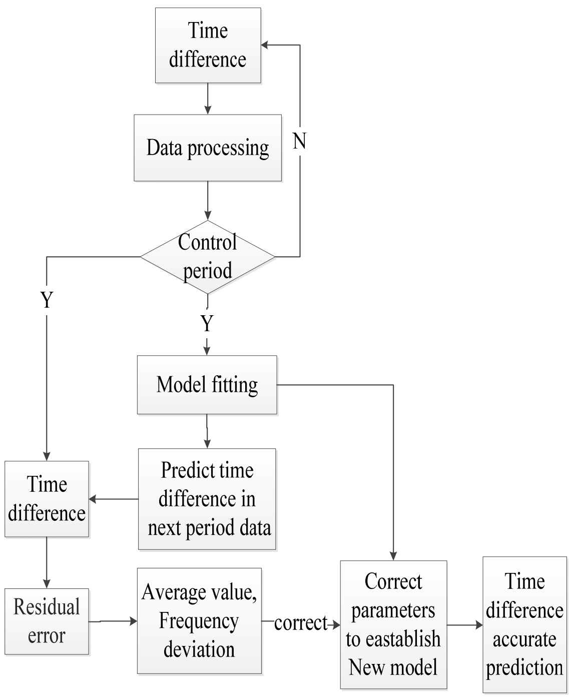


Fig. 1. The flow chart of residual error correction method.

Firstly, the time difference should be collected in real time. Until in the control period, the fitting method like least squares method should be used to establish the time difference model and calculate parameters such as phase, frequency deviation and frequency drift. Secondly, the predicted time difference in the next control period will be calculated by time difference model. Thirdly, when the true time difference of next control period is collected, the residual value can be obtained by the true value minus the predicted value. At last, it can calculate the average of the residual values of the control period in order to convert to the frequency deviation. The formula is as follows:

$$a_{\hat{\varepsilon}_1} = \frac{\sum_{j=1}^{t} \hat{\varepsilon}_j}{t} \quad (1)$$

In the formula, t represents the control period. By fitting the true value in the current control period to obtain $a_1(i+1)$. Then $a_1(i)$ is corrected with to make a new parameter $a_1'(i+1)$ and update a new time difference model. The formula is:

$$a_{1_{i+1}}^{'} = a_{1_i} - a_{\hat{\varepsilon}_i}, i = 1, 2...n \quad (2)$$

where i is the number of control period, $a_1(i+1)$ and $a_1'(i+1)$ are the first-order coefficient and the modified first-order coefficient of time difference model in i+1th control period respectively, $a_1(i)$ is the first-order coefficient of time difference model in i^{th} control period.

3.2. Dynamic Steering Method

The key idea of dynamic steering method of the atomic clock is change adjustment and control strategy to synchronize with the reference, which is according to the time difference between the atomic clock and the reference.

From the atomic clock difference model, it is obvious that the phase deviation and frequency deviation of the signal occur with timing. Based on test requirements and time difference data, adjustments and control strategies can be determined.

To determine the frequency adjustment, it is necessary to combine the time difference of the current time and the frequency deviation, which are mainly divided into four cases. It is as shown in Fig. 2.

In the first situation, it indicates that the atomic clock signal is deviating from the reference signal and the deviation speed is increased. In the second situation, the atomic clock signal is reversely deviated from the reference signal, and the deviation speed is increased.

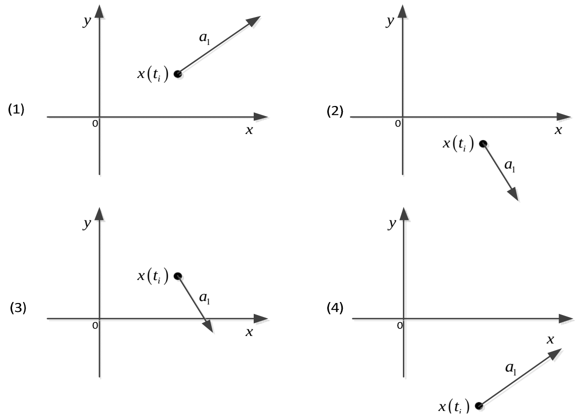


Fig. 2. Distribution of time difference and frequency deviation.

In the third case, it indicates that the atomic clock signal is approaching the reference signal from the positive direction, and the approach speed is increased. In the fourth case, it indicates that the atomic clock signal is close to the reference signal from the negative direction, and the approach speed increases.

If the first or second situation occurs, the frequency adjustment should be added to the current frequency deviation. If the third or fourth situation occurs, the frequency adjustment should be subtracted from the current frequency deviation, or the frequency adjustment should not be changed to waiting for infinite proximity.

According to the status of reference source, there are two strategies for steering which are adjustment in real time and autonomous maintenance.

When the reference source is normal, the time difference is collected and predicted in real time. according to historical data, and the trend of time difference is judged. The steering method adopts minimum control adjustment strategy, combined with the atomic clock performance, to adjust the controlled atomic clock slowly. The method maintains the accuracy and frequency stability of the output signal from controlled atomic clock. The flow chart is shown in Fig. 3.

When the time difference is great than threshold, phase adjustment is used firstly to make the phase near zero. Otherwise, time difference model is fitted by correcting the residual error. According to modified time difference model, the predicted time difference is calculated which comparing to another threshold. When the predicted time difference is great than threshold, the key parameters must be changed to adjust the clock.otherwise, the key parameters should keep for this time.

When the reference source is abnormal, previous time difference must be used. The flow chart is as shown in Fig. 4.

The time difference is predicted in real time according to historical data, and the trend of time difference is judged. When the time difference diverges, the adjustment is step by step to send to the controlled clock.

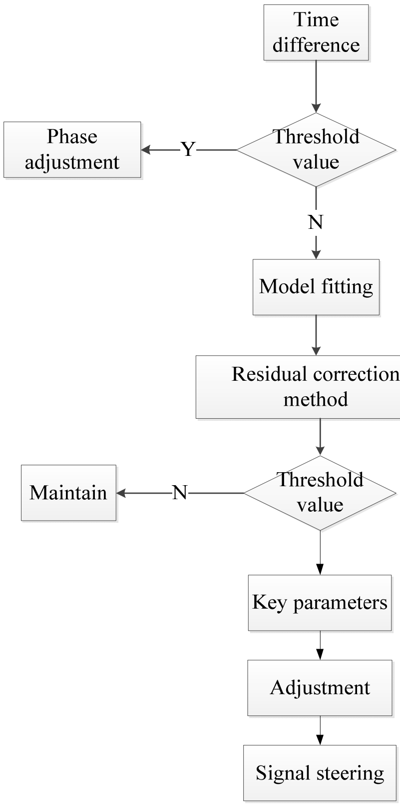


Fig. 3. Distribution of time difference and frequency deviation.

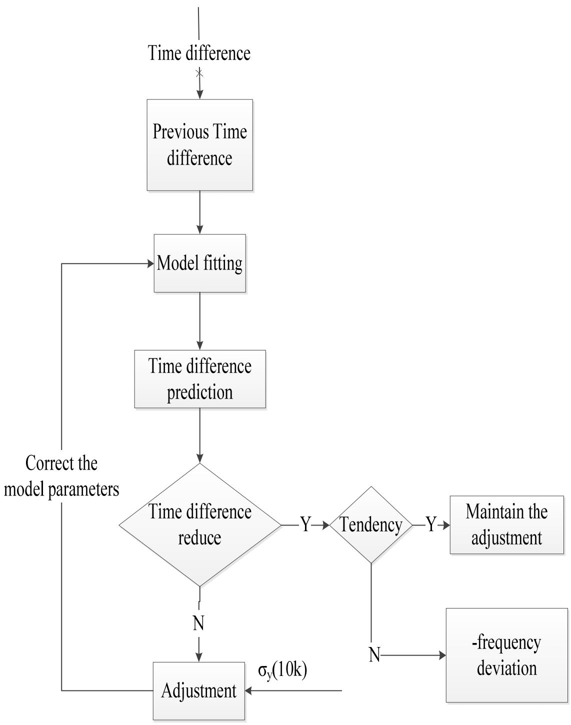


Fig. 4. Distribution of time difference and frequency deviation at abnormal reference source.

When the time difference converges, the adjustment for signal is calculated in real time based on the change trend. Until the tendency of predicted time difference reduces, the adjustment should be maintained. Otherwise, the opposite of frequency deviation must be sent to atomic clock.

4. Experimental Results

In order to further verify the dynamic control method of atomic clock based on residual correction, a test platform is designed and built. The hardware of the test platform is composed of atomic clock, high resolution offset generator(HROG-10), time interval counter(SR620), multi-channel phase measuring system(TSC MMS) and computer; the software is composed of atomic clock control software running on the industrial computer. The test platform structure is as shown in Fig. 5.

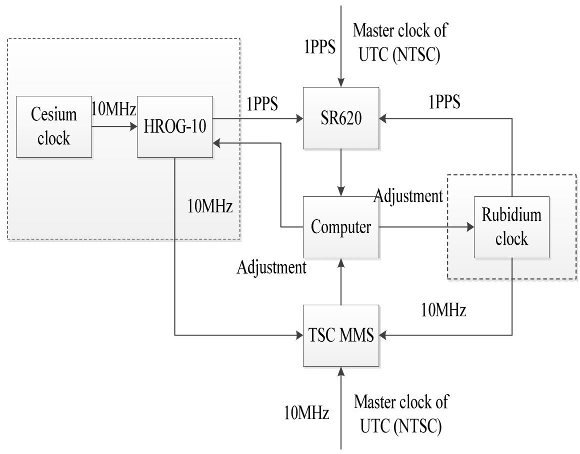


Fig. 5. The structure diagram of test platform for Cesium and Rubidium clock steering method.

Firstly, the output signal of atomic clock is divided in two signals which are 1PPS and 10 MHz signal through HROG-10. They are connected to SR620 and TSC MMS respectively. SR620 is responsible for measuring the time difference between two 1PPS signals outputted from HROG-10 and master clock of UTC (NTSC). TSC MMS is responsible for measuring the phase difference between two 10 MHz signals outputted from HROG-10 and master clock of UTC (NTSC). Secondly, the atomic clock control software designed collects the time difference and phase difference every second. According to the type of the controlled atomic clock, the key parameters and the control strategy are determined. Thirdly, the obtained adjustment is sent to HROG-10 or the atomic clock through the RS232 serial port. Therefore, it can make sure the time difference of the controlled clock signal within a certain range. At the same time, according to the phase difference of the two signals measured by TSC MMS, the frequency stability of the controlled atomic clock is calculated to determine whether the steering method affects the frequency stability of the controlled atomic clock output signal or not.

After the measurement platform is built, the master clock of UTC(NTSC) is used as the reference signal, and the output signals of Cesium clock and Rubidium are used as controlled clock respectively. The measurement results are shown in Figs. 6-11 and Table 1 that including the results of time difference and frequency stability.

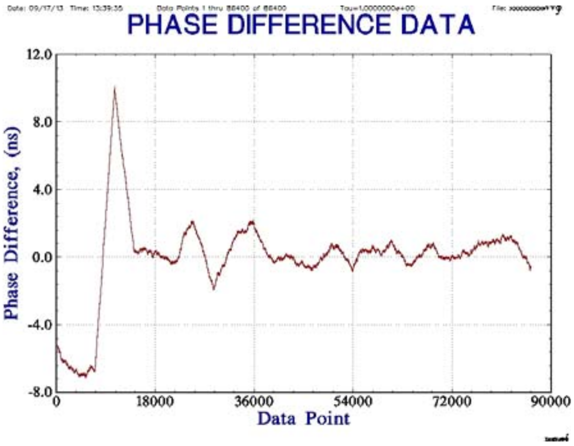


Fig. 6. The time difference data of Cesium clock steering for one day.

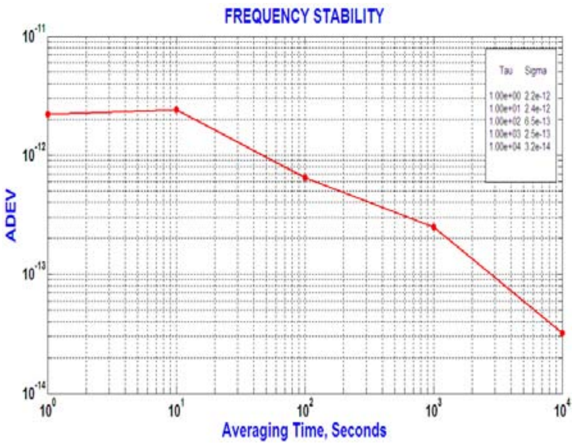


Fig. 7. Frequency stability with Cesium clock steering.

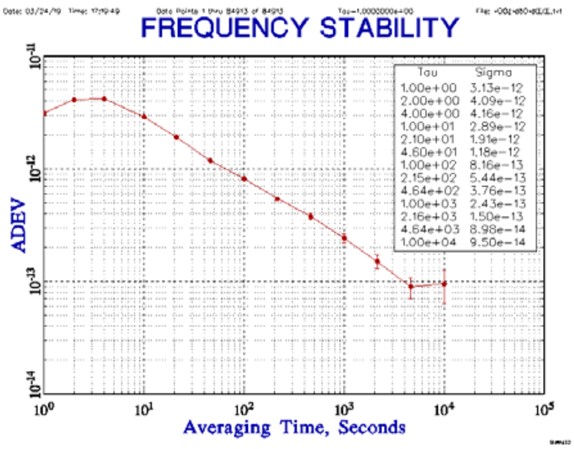


Fig. 8. Frequency stability without Cesium clock steering.

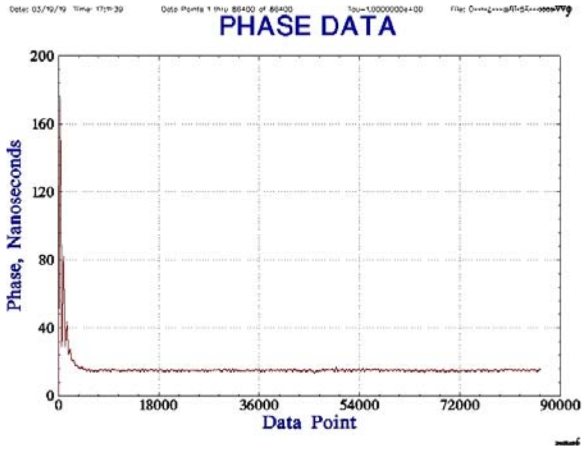


Fig. 9. The time difference data of Rubidium clock steering for one day

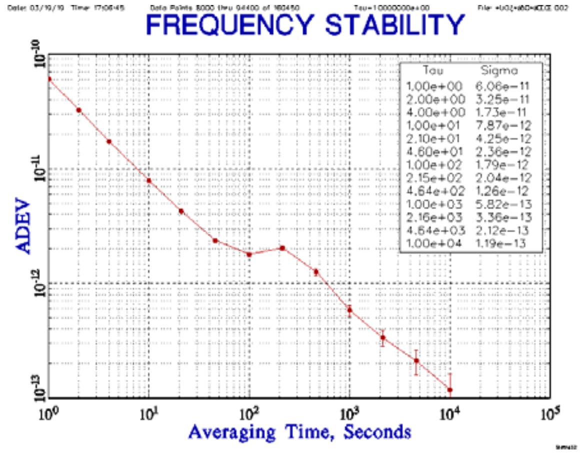


Fig. 10. Frequency stability with Rubidium clock steering.

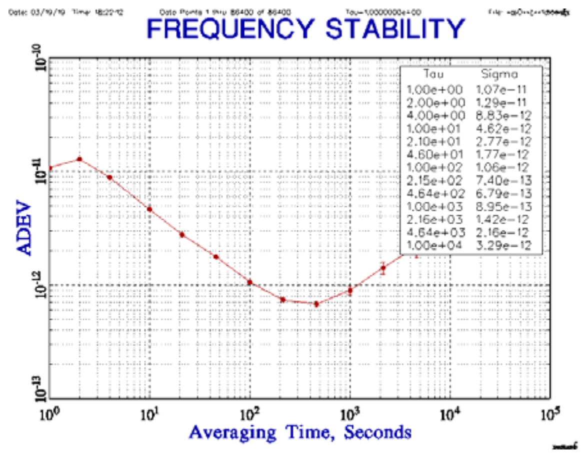


Fig. 11. Frequency stability without Rubidium clock steering.

It can be seen from Figs. 6-11 that the time difference of the running freely of Cesium clock for 16 hours is 9.4 ns. After using the dynamic steering method based on residual correction, the time difference of one day is within 2 ns. When the Rubidium clock is running freely, the time difference of one day changes quickly about 11 μs. After using the same steering method, the time difference of one day is less than 4 ns. The results show that the atomic clock dynamic steering method based on residual correction is feasible for atomic clock frequency control, especially for atomic clocks with lower frequency accuracy.

Table 1. Frequency stability of Rubidium clock and Cesium clock under two status.

Tau(s)	1	10	100	1000	10000
Cesium clock without steering	3.13 E-12	2.89 E-12	8.16 E-13	2.43 E-13	9.50 E-14
Cesium clock with steering	2.20 E-12	2.40 E-12	6.50 E-13	2.50 E-13	3.20 E-14
Rubidium clock without steering	1.07 E-11	4.62 E-12	1.06 E-12	8.95 E-13	3.29 E-12
Rubidium clock with steering	6.06 E-11	7.87 E-12	1.79 E-12	5.82 E-13	1.19 E-13

It also can be seen from Table 1 that the short-term frequency stability of the output signal of controlled Cesium clock is almost the same as which of the uncontrolled clock. Moreover, the long-term frequency stability of controlled Cesium clock is improved than which of uncontrolled clock. The short-term frequency stability of the controlled Rubidium clock is slightly lower than which of the uncontrolled clock, and the long-term frequency stability is significantly improved. The accuracy of the signal and long-term frequency stability is greatly improved with the proposed method.

The results show that the frequency stability is closely related to the measurement period and the control strategy. The measurement period of the Rubidium clock is much smaller than the period of Cesium clock. It is a very effective way to improve the long-term frequency stability of the signal by steering the atomic clock in real time.

5. Conclusion

In summary, the dynamic steering method proposed is useful for atomic clock. It makes full use of residual error to improve the accuracy of time difference prediction. The result of synchronization is at least twice as much as that of traditional methods.

The method can improve the accuracy and stability of the output signal for the atomic clock, and effectively improves the long-term stability of the signal, and shows an outstanding performance on atomic clocks with low frequency accuracy.

Acknowledgements

Thanks for the support of Youth innovation promotion association of Chinese Academy of Sciences (1180000XYR).

References

[1]. Guo Yuan, The Design of Time-frequency Signal Production and Maintenance System of Satellite, *XIDIAN University*, 2013.

[2]. Fan Duosheng, Shi Shaohua, Li Xiaohui, An Algorithm for the Rubidium Atomic Clock Control Based on the Kalman Filter, *Journal of Astronautics*, Vol. 1, Issue 36, 2015, pp. 90-95.

[3]. Li Yuwei, Wang Wenli, Ya Liu, Li Xiaohui, Dong Ruifang, Liu Yinhua, Study of an atomic clock steering method based on least square, in *Proceedings of the IEEE International Frequency Control Symposium*, Taipei, Taiwan, 19-22 May 2014.

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 42-48

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Route Towards a Label-free Optical Waveguide Sensing Platform Based on Lossy Mode Resonances

[1, 2, *] C. R. Zamarreño, [1] P. Zubiate, [1] A. Ozcariz, [1] C. Elosua, [1] A.B. Socorro, [1] A. Urrutia, [1] D. Lopez, [1] N. De Acha, [1] J. Ascorbe, [1] I. Vitoria, [1] J. J. Imas, [1] J. M. Corres, [1, 2] S. Diaz, [3] M. Hernaez, [1, 2] J. Goicoechea, [1, 2] F. J. Arregui, [1, 2] I. R. Matías and [1, 2] I. del Villar

[1] Institute of Smart Cities (ISC), Public University of Navarra, Pamplona, 31006, Spain
[2] Electrical, Electronic and Communication Engineering Department, Public University of Navarra, Pamplona (Navarra), 31006, Spain
[3] School of Chemistry, Faculty of Science, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom
[*] Tel.: + 34948168445, fax: + 34948169720
[*] E-mail: carlos.ruiz@unavarra.es

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: According to recent market studies of the North American company Allied Market Research, the field of photonic sensors is an emerging strategic field for the following years and it is expected to garner $18 billion by 2021. The integration of micro and nanofabrication technologies in the field of sensors has allowed the development of new technological concepts such as lab-on-a-chip which have achieved extraordinary advances in terms of detection and applicability, for example in the field of biosensors. This continuous development has allowed that equipment consisting of many complex devices that occupied a whole room a few years ago, at present it is possible to handle them in the palm of the hand; that formerly long duration processes are carried out in a matter of milliseconds and that a technology previously dedicated solely to military or scientific uses is available to the vast majority of consumers. The adequate combination of micro and nanostructured coatings with optical fiber sensors has permitted us to develop novel sensing technologies, such as the first experimental demonstration of lossy mode resonances (LMRs) for sensing applications, with more than one hundred citations and related publications in high rank journals and top conferences. In fact, fiber optic LMR-based devices have been proven as devices with one of the highest sensitivity for refractometric applications. Refractive index sensitivity is an indirect and simple indicator of how sensitive the device is to chemical and biological species, topic where this proposal is focused. Consequently, the utilization of these devices for chemical and biosensing applications is a clear opportunity that could open novel and interesting research lines and applications as well as simplify current analytical methodologies. As a result, on the basis of our previous experience with LMR based sensors to attain very high sensitivities, the objective of this paper is presenting the route for the development of label-free optical waveguide sensing platform based on LMRs that enable to explore the limits of this technology for bio-chemosensing applications.

Keywords: Lossy mode resonances, Optical fiber sensor, Thin-films, Sensing platform, Refractive index.

1. Introduction

Photonic sensor detection systems are attracting a great deal of interest and investment owing to their superior performance [1-3] and the utilization of intimate light-matter interactions is becoming more and more attractive for both chemical and biological detection of different species in size (from nm to microns) at surprisingly low concentrations (attomolar or parts per trillion or even further) [4] capable to cope with very challenging applications, such as cell activity or volatile compounds detection in aqueous or gaseous environments respectively [5]. So, there is a need for the development of cheap, portable and real time detection devices, such as the one represented schematically in Fig. 1, applicable in various fields like food quality, discovery of novel biomarkers and early diagnosis in medicine, the recognition of chemical substances at very low dilutions in chemical sciences, trace detection of air pollutants at very low concentrations associated to the more and more restrictive regulations in harmful gases emissions or other applications.

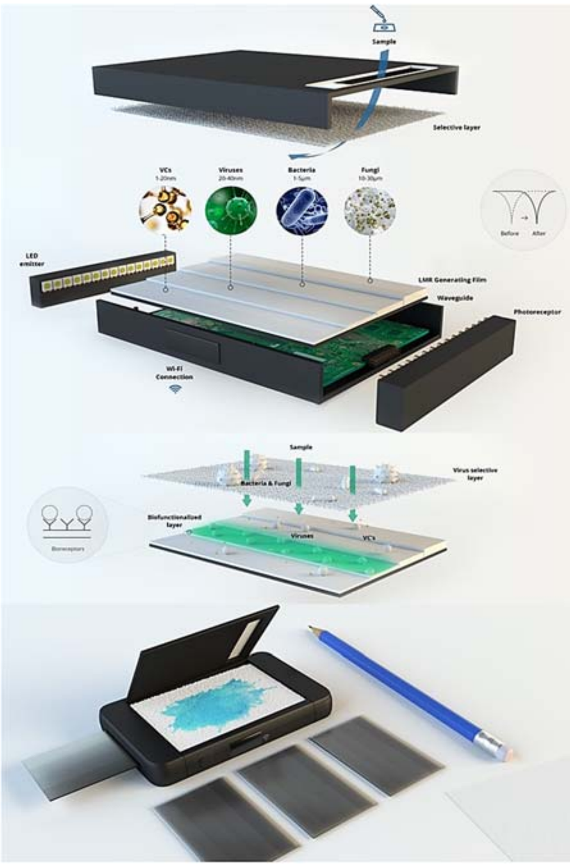


Fig. 1. Setup of the label-free multipurpose sensor. At the top, a schematic of the sensing platform. In the middle, a practical application of the platform towards detection of a type of virus. At the bottom, an infographic of the final device with the detection cartridges.

Currently, photonic sensing systems are far from being exploited to their full potential and there is still a long way to go in the understanding of these phenomena in order to achieve so. For example, the first SPR based sensor developed in 1982 was a scientific breakthrough [6]. It used the Kretschmann-Raether configuration [7], which basically consisted of an optical prism with a metallic overlay of tens of nanometres fabricated on one facet (see Fig. 1-top), which can be simplified using an optical fiber configuration (see Fig. 1-bottom) that removes the prism while reduces the size of the setup. Both setups presented in Fig. 2 permit surface plasmon polaritons to be generated at the metal-dielectric interfaces and to couple light at specific wavelength ranges as a function of the surrounding medium conditions. This configuration can be adjusted for biological or chemical applications adding a selective coating for the selected target on top of the metallic layer, then a biosensor or a chemical sensor is obtained [8].

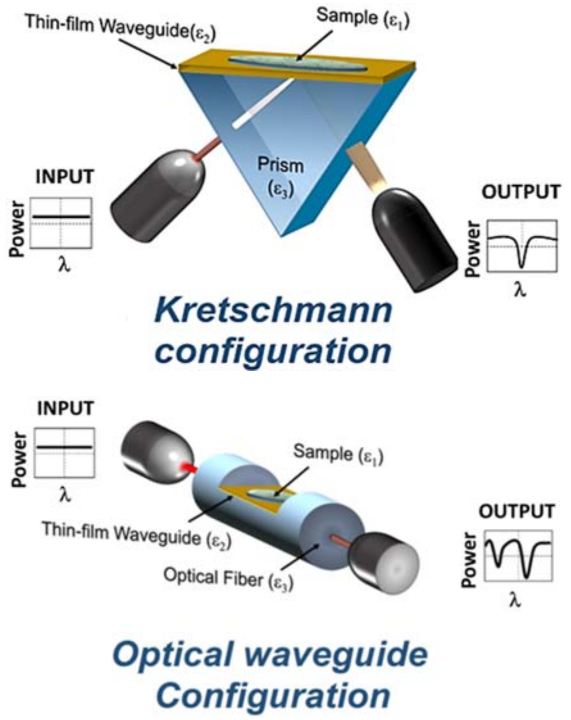


Fig. 2. Resonance generation with a nanocoated optical prism (SPR, Kretschmann configuration, top) and a nanocoated D-shaped fiber optic, LMR (bottom). Reprinted from [9] with permission from Elsevier.

Since SPR-based sensing devices working principle remains the same, the technological platforms have evolved in terms of stability, robustness and applicability with the possibility to detect different analytes at different concentrations at the same time. As a result, SPR based sensors are expected to reach in 2025 a business of more than 1 billion dollars [10].

The interest in this kind of devices has been already proven by commercial developments of SPR-based devices such as GE-healthcare Biacore systems for screening and quality characterization of molecular interactions from ions to viruses [11]. In the same manner, other devices based on different optical principles have emerged, such as Research International RAPTOR, a 4-channel optical waveguide system for rapid, automatic, portable fluorometric assay for monitoring toxins, viruses, bacteria, spores, fungi and other diverse targets [12] or Fortebio OCTET that performs quantitation and kinetics characterization assays at high speed and throughput [13]. However, previous devices are mainly intended for analysis in aqueous environments, the size of the equipment prevents in most cases from any portable measurement or in-field usage and they do not enable analysis in gaseous media, which is an important question that must be addressed for a versatile analytical tool.

The solution to previous concerns might come by exploiting lossy mode resonances (LMR), a phenomenon that emerges from the promising combination of nanotechnology (a thin-film with nanometer thickness deposited onto a substrate) and photonics (light transmitted through a waveguide is coupled to the thin-film generating a resonance).

The first reported experimental LMR demonstration was performed by Dr. Carlos Ruiz Zamarreño in his PhD thesis followed by a complete theoretical explanation of the phenomenon by Dr. Ignacio Del Villar in [14], the corresponding diffusion by Professor Ignacio Raúl Matias in an invited talk at Optical Fiber Sensors Conference 2009 followed by the invited talk given by Dr. Carlos Ruiz Zamarreño at SEIA conference in 2019 [15]. During that time, this technology has been cited more than 3,000 times in different projects related to LMR-based devices for the detection of pH [9], VCs [16], IgGs [17], proteins [18] or DNA [19] just to mention a few. In addition, other research groups from United Kingdom, Poland, India or China have explored the potential of LMRs for sensing applications in both experimental and theoretical works [20-27].

LMR technique has been proposed as an alternative sensing platform that competes with SPR. Both phenomena occur when the magnitude of the real part of the thin-film permittivity is greater than both its imaginary part and the permittivity of the material surrounding the thin-film (dielectric medium). However, unlike in SPRs, where the real part of the thin-film permittivity is positive (that is, metallic material) in LMRs it is negative (dielectric or semiconductor material) [28]. As a result, LMR based sensors broaden the range of materials that allow attaining a resonance and the associated chemistry possibilities.

Moreover, LMRs overcome SPRs in optical waveguide configurations (typically fiber optic where the incident angle is close to 90º), where the resonances are induced when a mode guided in the waveguide experiences a transition to guidance in a thin-film with losses. LMRs offer other important differences in the way they operate, which is advantageous for sensing purposes: LMRs can be obtained with both transverse electric (TE) and transverse magnetic (TM) polarized light while SPRs are only obtained with TM polarized light [29], which permits to obtain reference measurements as well as override the utilization of polarized light if necessary. Moreover, the position of the resonance in the optical spectrum can be precisely tuned in a simple way by controlling the thin-film thickness (tens to hundreds of nm thick). An additional way to tune the LMR is through the dielectric properties of the thin-film and the optical waveguide as it is described in [30].

Among the LMR potential application domains, probably the most important ones, from the point of view of their multiplicity of possibilities, are volatile compound detection and biosensing. On one side, the detection of VCs in air could enable detecting explosives and distinguishing specific biomarkers in human breath or other body fluids for volatolomic applications towards early diagnosis and detection of acute pathologies and data mining [31]. On the biosensing side, the rapid and real time detection of DNA/RNA towards DNA methylation and DNA mutation identification to viruses, proteins and cell activity monitoring are of extreme importance.

2. Previous Results

Obtained results in this field rely on one of the key parameters of LMRs, "the sensitivity", which can be improved in a very simple way by controlling the thin-film thickness, the thin-film refractive index and the surrounding medium refractive index. These basic rules, described in one of our works [30], have allowed us to obtain a major advance in LMR with a record sensitivity below 10^{-9} refractive index units (RIU) per nanometer [32] (see Fig. 2), which permits for example to detect concentrations in the order of femtomolar or even lower, [17]. This record sensitivity was achieved in the vicinity of the refractive index of the substrate (~1.45) because, as stated in [30], the sensitivity increases exponentially as the surrounding medium refractive index approaches the refractive index of the substrate. Therefore, the adequate matching between the dielectric properties of the substrate and the surrounding medium as well as the optimization of the LMR generating layer or layers in terms of thickness and dielectric permittivity might open a full range of possibilities in the improvement and applications of these devices throughout an intensive study of material properties, fabrication procedures and technology integration.

In the case of Fig. 3, the substrate was optical fiber made of fused silica (refractive index~1.45). However, as stated before, the device has not been optimized for operation in the aqueous domain with a surrounding medium refractive index ~1.33 (it has applications in oil industry though [33]), which still presents an opportunity for a sensitivity improvement of several orders of magnitude. Concerning the application of these devices for biosensing applications the two most interesting regions are the refractive index region of water-based solutions (~1.33) and the refractive index region of air (~1) for aqueous and gaseous media respectively. Most of the chemical and biological reactions take place in aqueous media, whereas many other applications can be found in air, (detecting VCs for example). Therefore, from [30], the development of optical sensors using the adequate LMR generating thin-films and waveguide substrates with a refractive index of 1.33 or 1 is expected to satisfy the conditions for maximum sensitivity, which is not trivial to come through or straightforward to prove experimentally due to the difficulties to design or handle such a thin-films as well as to find suitable waveguide substrates with that properties towards the fabrication of a multipurpose label-free LMR-based platform.

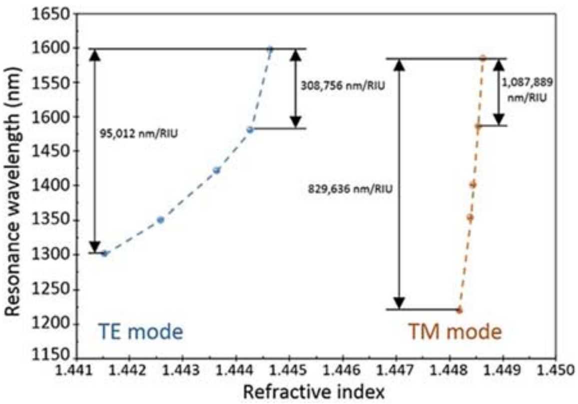


Fig. 3. LMR wavelength shift as a function of the surrounding medium refractive index using a D-shape fiber coated with a SnO2 thin-film. Reprinted from [32] with permission from Scientific Reports-Nature.

3. Discussion

The idea behind a label-free LMR-based sensing platform basically consist of exploiting the LMR phenomenon towards a wide range of applications in a simple and versatile manner by means of the combination of thin-film multi-pattern structures of different materials.

In particular, thin-film material dielectric properties enable to position the resonance within a defined working range that satisfies LMR generation conditions among other things [30]. Concerning the thin-film thickness, thicker coatings will excite LMRs at longer wavelengths and they will be more adequate for the detection of bigger size targets in the NIR spectrum, whereas thinner coatings will excite LMRs at shorter wavelengths and will be more adequate for the detection of small size targets in the UV spectrum. Fig. 4 gives an idea of the corresponding detection wavelength and thickness of the nanocoating for each target size (VCs, DNA, virus or cells).

Moreover, thin-film patterning permit to enhance the sensitivity and selectivity capabilities of the devices, which are mainly associated to the penetration depth of the evanescent wave (shorter for UV and larger for NIR) and the surface area. LMR properties can be predicted roughly using material theoretical models and waveguide propagation theory. However, there is little work done in this field, the models are not complete and often do not cover a broad spectral range (UV-VIS-NIR) or complex patterns, which requires a continuous experimental feedback to adjust correctly these models in order to obtain accurate calculations for the fabrication of LMR-based devices matched with the application.

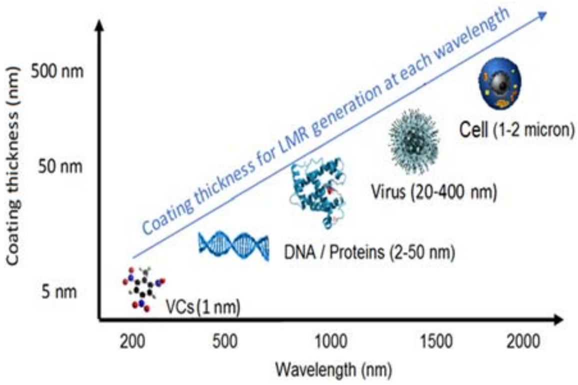


Fig. 4. Wavelength versus adequate coating thickness for detecting different targets: VCs, DNA, Proteins, viruses and cells.

Finally, the adequate selection of the optical waveguide geometry and material (the refractive index of the optical waveguide should be similar to that of the surrounding medium as it was mentioned before) and interrogation method enable to improve the signal to noise ratio and increase the resolution of the device. Here, the utilization of transmission and reflection configurations in both cylindrical and planar waveguides should be evaluated in order to opt for what it is called as “lab-on-fiber”, “lab-in-fiber” [34] or “planar waveguide” [35] configurations (see Fig. 5).

Fig. 6 shows some simulation results for a substrate of refractive index 1.45 (silica) where a coating of refractive index 2, like that of SnO_2 [32], has been used as LMR generating thin-film. In a simple manner, Fig. 6 shows the LMR relative wavelength shift as a function of the thickness of the LMR generating thin-film when a second coating representing the size of VCs (5 nm), viruses (20 nm) or cells (100 nm) are attached to the thin-film. Depending on whether the surface is covered by any of them, the resonance wavelength associated to each coating thickness will experience a different relative wavelength shift.

As it is depicted in Fig. 6, small molecules, such as VCs (in the order of a few nanometers) will lead to a measurable relative wavelength shift for thin LMR generating coatings. However, for thick LMR generating coatings the resonance wavelength shift will be very low for VCs. In view that the absolute wavelength resolution is smaller at longer wavelengths, the effect of VCs will not be detected at long wavelengths, where LMRs induced by thick coatings are located. On the opposite side, the relative wavelength shift induced by thick biolayers (cells) when a thin LMR generating coating is used, exceeds 100 % and it is difficult to track, whereas for thicker LMR generating coatings this relative wavelength shift is smaller and adequate for the absolute resolution of the equipment. In other words, thin films are more appropriated for small feature detection while thick films are appropriated for larger size targets. All in all, the position of the resonance also takes an important role and is dependent on both the refractive index of the substrate (silica in the simulation) and the surrounding medium as well as the LMR generating material dielectric properties, which requires a profound knowledge of the application and the final performance of the device for a proper design.

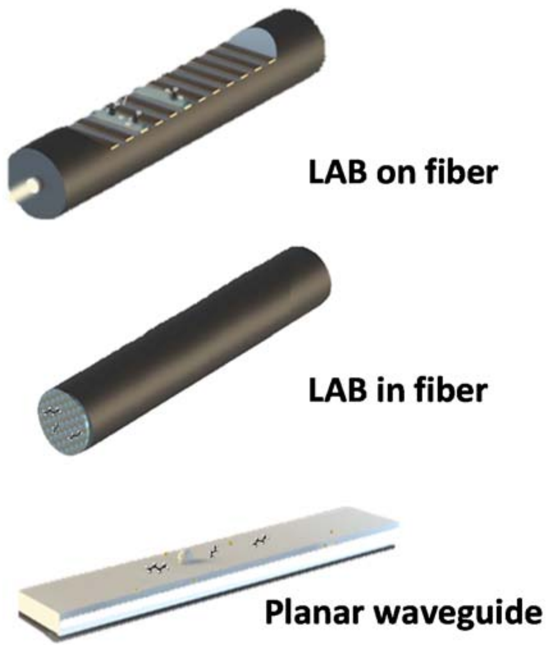


Fig. 5. Schematic representation of proposed optical waveguide configurations: lab on fiber (top), lab in fiber (center) and planar waveguide (bottom).

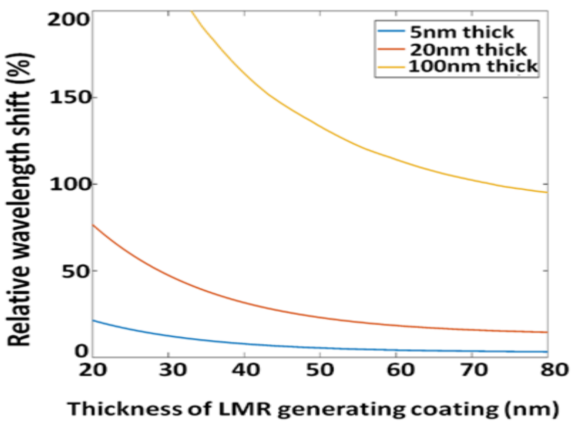


Fig. 6. LMR relative wavelength shift as a function of the coating thickness when a second coating representing the size of VCs (5 nm), viruses (20 nm) and cells (100 nm) is attached.

4. Challenges

It is important to remark that the development of the label-free LMR-based sensing platform proposed in Fig. 1 should also overcome many drawbacks concerning the utilization of broadband excitation light sources (covering the entire working range). At the same time, the light coupled through the optical waveguides should be collected by the corresponding detectors in the entire working range. In addition, the profile (thickness) and pattern of the device should be chosen according to the desired target size (from nm to microns), which might require the utilization of a porous membrane or filter that prevent particles that exceed the objective target size to be detected and avoid interferences and undesirable resonance shifts produced by bigger particles attached to thinner resonance generating thin-films as it was mentioned before when discussing the simulations of Figure 6. Finally, the chemistry of the sensitive layers used as receptors should be compatible with the LMR generating material in order to form strong bonds that enable an effective target particle capture.

5. Outlook

The achievement of this ambitious purpose could open infinite possibilities in science and knowledge generation, such as early diagnoses that are not possible nowadays either for the low concentration of target analytes, such as VCs and gases (volatolomics) in breath or proteins in urine for example or for the necessity to perform target detection in air and liquid at the same time. Additionally, the utilization of the proposed system could be helpful for the discovery of unknown biomarkers in other body fluids, such as tears, skin secretions or flatulencies that were not possible to detect before, due to their low concentration or the difficulties in the application of standard detection methods. Health and safety are some of the most important sectors that could benefit from these achievements.

Acknowledgements

This work was supported in part by the Spanish Ministry of Education and Science-FEDER TEC2016-78047-R, Government of Navarra 0011-1365-2017-000117 and Public University of Navarra PJUPNA26 research grants as well as the ATTRACT call financed by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777222.

References

[1]. Preeti Bisht, Photonic Sensors Market - Sales Forecasts, Market Size, Share, Growth and Analysis, *Allied Market Research*, 2016.

[2]. Editorial, A small world full of opportunities, *Nature Materials*, Vol. 9, Issue 3, 2010, pp. 181-279.

[3]. Editorial, The hidden face of nanophotonics, *Nature Photonics*, Vol. 5, Issue 7, 2011, p. 379.

[4]. C. Caucheteur, T. Guo, F. Liu, B.-O. Guan, J. Albert, Ultrasensitive plasmonic sensing in air using optical fibre spectral combs, *Nature Communications*, Vol. 7, 2016, 13371.

[5]. J. Morla-Folch, P. Gisbert-Quilis, M. Masetti, E. Garcia-Rico, R. R. Alvarez-Puebla, L. Guerrini, Conformational SERS Classification of K-Ras Point Mutations for Cancer Diagnostics, *Angewandte Chemie - International Edition*, Vol. 56, Issue 9, 2017, pp. 2381-2385.

[6]. T. Nylander, C. Liedberg, B. Lind, Gas detection by means of surface plasmon resonance, *Sens. & Actuators*, Vol. 3, Issue 1, 1982, pp. 79–88.

[7]. H. Kretschmann, E. Raether, Radiative decay of nonradiative surface plasmons excited by light, *Zeitschrift für Naturforschung A*, Vol. 23, Issue 12, 1968, pp. 2135–2136.

[8]. J. Homola, Surface plasmon resonance sensors for detection of chemical and biological species, *Chem. Rev*., Vol. 108, Issue 2, 2008, pp. 462-493.

[9]. I. del Villar, F. J. Arregui, C. R. Zamarreño, *et al*., Optical sensors based on lossy-mode resonances, *Sens. & Actuators B*, Vol. 240, 2017, pp. 174-185.

[10]. F. M. Insights, Surface Plasmon Resonance (SPR) Market to be worth US$ 1,110.4 Mn by the end of 2025, *Future Market Insights Global and Consulting Pvt. Ltd.,* 2017.

[11]. J. A. Lofgren, *et al.*, Comparing ELISA and surface plasmon resonance for assessing clinical immunogenicity of panitumumab, *Journal of Immunology*, Vol. 178, Issue 11, 2007, pp. 7467-7472.

[12]. K. D. King, J. M. Vanniere, J. L. Leblanc, K. E. Bullock, G. P. Anderson, Automated fiber optic biosensor for multiplexed immunoassays, *Environmental Science and Technology*, Vol. 34, Issue 13, 200, pp. 2845-2850.

[13]. Y. Abdiche, D. Malashock, A. Pinkerton, J. Pons, Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet, *Analytical Biochemistry*, Vol. 377, Issue 2, 2008, pp. 209-217.

[14]. I. Del Villar, C. R. Zamarreño, M. Hernaez, F. J. Arregui, I. R. Matias, Lossy mode resonance generation with indium-tin-oxide coated optical fibers for sensing applications, *J. Light. Technol*., Vol. 28, Issue 1, 2010, pp. 111–117.

[15]. C. R. Zamarreño, P. Zubiate, A. Ozcariz, C Elosua, *et al.*, Label-free optical fiber sensing platform based on Lossy Mode Resonances, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'2019)*, Tenerife (Canary Islands), Spain, 25-27 September 2019, pp. 130-131.

[16]. C. Elosua, F. J. Arregui, C. R. Zamarreño, *et al*., Volatile organic compounds optical fiber sensor based on lossy mode resonances, *Sens. & Actuators B*, Vol. 173, 2012, pp. 523-529.

[17]. F. Chiavaioli, P. Zubiate, I. del Villar, *et al*., Femtomolar Detection by Nanocoated Fiber Label-Free Biosensors, *ACS Sensors*, Vol. 3, Issue 5, 2018, pp. 936-943.

[18]. P. Zubiate, C. R. Zamarreño, P. Sánchez, I. R. Matias, F. J. Arregui, High sensitive and selective C-reactive protein detection by means of lossy mode resonance based optical fiber devices, *Biosens. Bioelectron*., Vol. 93, 2017, pp. 176–181.

[19]. C. R. Zamarreño, P. Zubiate, P. Sanchez, *et al*., Single strand DNA detection by means of lossy mode resonance-based optical fiber devices, in *Proceedings of the IEEE Sensors*, 2016, 7808621.

[20]. D. Tiwari, K. Mullaney, S. Korposh, S. W. James, S. W. Lee, R. P. Tatam, An ammonia sensor based on Lossy Mode Resonances on a tapered optical fibre coated with porphyrin incorporated titanium dioxide, *Sens. Act. B: Chem*, Vol. 242, 2017, pp. 645-652.

[21]. Miguel Hernaez, Andrew G. Mayes, Sonia Melendi-Espina, Graphene oxide in lossy mode resonance-based optical fiber sensors for ethanol detection, *Sensors*, Vol. 18, Issue 1, 2018.

[22]. M Śmietana, *et al*., Optical monitoring of electrochemical processes with ITO-based lossy-mode resonance optical fiber sensor applied as an electrode, *J. Lightwave Technol*., Vol. 36, Issue 4, 2018, pp. 954-960.

[23]. S. P. Usha, A. M. Shrivastav, B. D. Gupta, Semiconductor metal oxide/polymer based fiber optic lossy mode resonance sensors: A contemporary study, *Optical Fiber Technol*., Vol. 45, 2018, pp. 146-166.

[24]. Nidhi Paliwal, Joseph John, Lossy Mode Resonance (LMR) Based Fiber Optic Sensors, *IEEE Sens. J*., Vol. 15, Issue 10, 2015, pp. 5361-5371.

[25]. D. Kaur, V. K Sharma, A. Kapoor, High sensitivity lossy mode resonance sensors, *Sens. Act. B: Chem*, Vol. 198, 2014, pp. 366–376.

[26]. L. Wu, *et al*., High-Performance Lossy-Mode Resonance Sensor Based on Few-Layer Black Phosphorus, *J. Phys. Chem. C*, Vol. 122, Issue 13, 2018, pp. 7368–7373.

[27]. S. P. Usha, A. M. Shrivastav, B. D. Gupta, Semiconductor metal oxide/polymer based fiber optic lossy mode resonance sensors: A contemporary study, *Optical Fiber Technol*., Vol. 45, 2018, pp. 146-166.

[28]. F. Yang, J. R. Sambles, Determination of the optical permittivity and thickness of absorbing films using long range modes, *J. Mod. Opt*., Vol. 44, Issue 6, 1997, pp. 1155–1164.

[29]. C. R. Zamarreño, M. Hernáez, I. Del Villar, I. R. Matías, F. J. Arregui, Optical fiber pH sensor based on lossy-mode resonances by means of thin polymeric coatings, *Sens. & Actuators B*, Vol. 155, Issue 1, 2011, pp. 290–297.

[30]. I. Del Villar, *et al*., Design rules for lossy mode resonance-based sensors, *App. Opt*., Vol. 51, Issue 19, 2012, pp. 4298–4307.

[31]. A. Mashir, R. A. Dweik, Exhaled breath analysis: The new interface between medicine and engineering, *Adv. Powder Technol*., Vol. 20, Issue 5, 2009, pp. 420–425.

[32]. A. Ozcariz, C. R. Zamarreño, P. Zubiate, F. J. Arregui, Is there a frontier in sensitivity with Lossy mode resonance (LMR) based refractometers?, *Scientific Reports*, Vol. 7, Issue 1, 2017, 10280.

[33]. P. Sánchez, C. R. Zamarreño, F. J. Arregui, I. R. Matías, LMR-based optical fiber refractometers for oil degradation sensing applications in synthetic lubricant oils, *Journal of Lightwave Technology*, Vol. 34, Issue 19, 2016, pp. 4537-4542.

[34]. M. Consales, *et al*., Lab-on-fiber technology: Toward multifunctional optical nanoprobes, *ACS Nano*, Vol. 6, Issue 4, 2012, pp. 3163-3170.

[35]. O. Fuentes, I. del Villar, J. M. Corres, I. R. Matias, Lossy mode resonance sensors based on lateral light incidence in nanocoated planar waveguides, *Scientific Reports*, Vol. 9, Issue 1, 2019, 8882.

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

A Loosely Coupled Sensing System Architecture and Implementation for Industrial IoT

*** Yasutaka SERIZAWA and Yusuke SHOMURA**
Hitachi America Ltd. R&D, IoT Edge Laboratory,
2535 Augustine Drive, 3rd Floor, Santa Clara, CA 95054, USA
Tel.: +1-408-986-6300
E-mail: yasutaka.serizawa@hal.hitachi.com

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: Sensing data increasingly become valuable source of information that enables various IIoT (Industrial IoT) applications. Traditionally, especially in manufacturing field sites, sensing data acquired by kinds of sensors via PLCs have been stored on storages with kinds of file formats, which has also been consumed by batch-type analytical processes as data sources. This system architecture has been one of the causes of lowering production efficiency such as long unplanned downtime of production machines. Taking this situation into account, we have proposed and designed a loosely coupled sensing system architecture, which is followed by some trial implementations into industrial IoT field. Acoustic sensors were applied as the first implementation of scalable and reusable sensing system, which supports streaming data processing capability. The developed system has pluggable system architecture where data acquisition part, signal processing part, and analytics part can be mutually communicated with common APIs, which provides flexible combination of streaming data processing in multiple parts. The scalability of developed system was demonstrated on two example applications of machine state classification and anomaly detection. The successful operation of both applications on the same acoustic sensing system indicates that the loosely coupled / pluggable system architecture works as expected, which suggests the developed system can reduce time and cost for applications establishment significantly.

Keywords: Loosely coupled sensing system, Real-time streaming data processing, Signal processing, Data acquisition, Acoustic analytics, One-stop configuration.

1. Introduction

The market of IIoT especially in manufacturing field has been growing increasingly. Given this trend, recently many manufacturers have become interested in whole process optimization using IT technologies. Therefore, introduction of IoT system has been widely considered especially in manufacturing factory shop floor in order to improve productivity and increase added value of products. However, it has been found that fully organized information (data) doesn't exist in existing manufacturing systems to monitor total condition or status of overall manufacturing shop floor in real-time manner, which provides lack of insights into shop floor operations.

Furthermore, there are many unscheduled downtimes which significantly reduce productivity not only due to sudden stops of manufacturing machinery or lines, but also lack of operators' awareness of wrong situations [1-2]. One of the root causes of unscheduled downtime would be lack of real-time and effective feedback to field site. From IoT system point of view, file-based data storage and on-demand data analysis which a lot of factories still use would provide this problem. They tend to try to execute some analysis and come up with some countermeasures after wrong situations happen.

Another critical problem which many manufacturers become to face on is aging workforce. As baby boomers reach retirement, the manufacturing industry is facing a skills dilemma [3]. Aging population impacts not only available labor force but also overall production quality, because aging and skilled workers in manufacturing field site play very important roles in keeping manufacturing processes high-quality one [4]. Skilled workers use their accumulated knowhow including human 5 senses, which is crucial to ‘feel’ and identify the situation of manufacturing, impending issues, and their root causes. Therefore, many manufacturers are struggling on the way to replace the skilled workforce, because training new employees is expensive and a drawn-out process. Furthermore, difficulty in replacement of accumulated knowhow and it’s positive impact on overall production processes bothers them.

According to this situation around manufacturing IoT, we have focused on real-time feedback to field site and additional-sensors approach which can be replaced by skilled workers’ accumulated knowhow based on 5 senses. Additionally, sensing system should be scalable and reusable in order to be applied to multiple environments seamlessly.

In this background, our research direction has focused on loosely coupled sensing system which supports streaming data processing with pluggable APIs and one-stop system configuration. Many works on sophisticated analytics have been disclosed and provided kinds of applications (e.g. [5-6]). However, almost all of them provide sophisticated analytics especially optimized to solve each specific problem. They also provide only analytics part from sensing system view point. In order to make sensing system flexible, or scalable, it should have pluggable APIs between function blocks as well as function modules, and configurable functionality from analytics part which can be strongly connected to applications. In this paper, the basic analytics-centric architecture of a loosely coupled sensing system which provides capability of streaming data process, pluggable APIs between function modules, and one-stop system configuration, as well as verification results for its scalability with applying to example applications, are reported.

2. A Loosely Coupled Sensing System

2.1. Data Processing System in Existing Manufacturing Site

In many cases of IIoT, especially in manufacturing field site, there are still many static systems for data analysis utilized. Acquired data tend to be stored on databases, or any formats of files (e.g. xml, xls, etc.) and are processed when problems occur or are checked by manufacturing operators periodically (e.g. daily, or weekly), which leads to long unplanned downtime, inefficient ways of manufacturing operation, and unoptimized assignment of resources.

Taking this situation as a general problem in manufacturing field site, we have proposed streaming data processing system enabling real-time data processing and having analytics-centric APIs especially in acoustic sensing field [7]. This also enables rapid establishment of applications especially on manufacturing use cases (scalability for application establishment). Fig. 1 shows conventional system of data processing in typical manufacturing field site, where data are basically generated by PLCs, and stored and utilized by processing with file-based media. In addition, when applying sophisticated analytics solutions to this kind of typical data processing system in manufacturing field site, batch-type data processing scheme would be applied, and configuration interfaces would be usually located on each part of data processing (data acquisition, signal processing, and analytics), which makes the system complicated and difficult to manage. In Section 2.2, our proposed sensing system architecture and application to acoustic sensing implementation including the difference from conventional scheme is described.

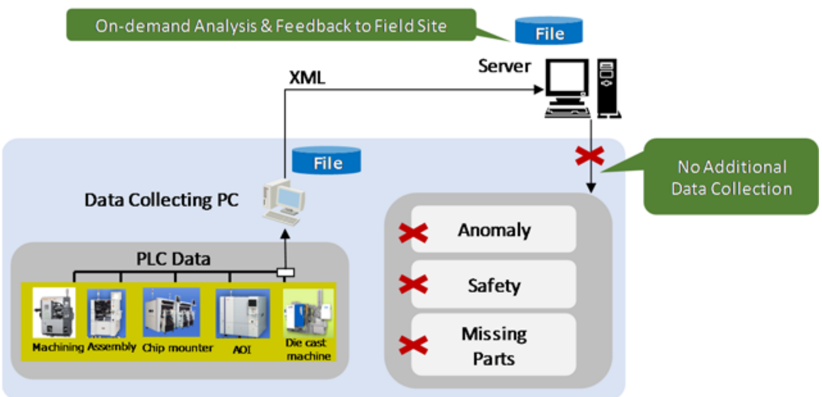


Fig. 1. Typical data processing system in manufacturing field site.

2.2. Proposed Sensing System Architecture and an Application to Acoustic Sensing

Fig. 2 (left) shows proposed overall system architecture. The streaming data generated from frequency sensors (vibration sensor, acoustic sensor, ultra-sonic sensor, and current sensor) in data acquisition part should be consumed by loosely coupled function modules which has well-defined API in signal processing part, and also transferred to analytics function modules in real-time streaming manner. The sensing system should have unified way to handle all the data sources (frequency sensors as well as file-based data), which requires certain adaptation layer in data acquisition part to extract kinds of sensors and provide API to signal processing part. The other requirement for the sensing system should be one-stop configuration from analytics side. Sensing application requires many parameters such as sampling frequency, streaming size (chunk size), bit width, filter types, filter coefficients, and so on. As described above, distributed allocation of these configuration information which conventional data processing systems adopt makes the system complicated and difficult to scale. In order to avoid such situation, our proposed system has directed to the one-stop configuration architecture.

Based on this overall system architecture, the first implementation of the system had focused on acoustic sensors. We chose acoustic sensors as the first target for the reasons of (1) acoustic sensors are promising data source for initial symptom detection of wrong situation in field sites, (2) the intrinsic features of 'contactless', 'scalable', and 'site agnostic', which are quite important from a view point of easy application to multiple fields, and (3) they are intuitive, which provides easy understanding of the applications for workers. Fig. 2 (right) shows detail description of an implementation of our acoustic streaming data processing system. In this system, gRPC [8] has been utilized to reduce message size and support scalability in terms of programming languages. One of the features of gRPC is characterized by serialization of data using protocol buffers instead of JSON and its high-speed communication capability (in other words, reduction of message size). Utilizing protocol buffers also provides fixed data structure designated by system, which can specify the data structure to other system as an API. In our implementation of acoustic streaming data processing system, protocol buffer provides an API of data acquisition part and data processing part to analytics part, which means that this system is analytics-centric design. gRPC server and client are deployed on each part (analytics, signal processing, and data acquisition), which enables flexible communication between individual parts. Within gRPC, request from client and response from server are exchanged based on protocol buffer as shown in Fig. 3. This structure enables flexible combination of signal processing and data acquisition from analytics point of view. Additionally, owing to broker functionality on signal processing part and data acquisition part, analytics can choose any kind of data from any kind of data source (e.g. data from device (raw streaming data), processed data with designated signal processing, or files).

In Fig. 4, a simple example implementation of proposed acoustic sensing system which has been implemented in some use cases described in chapter 3 is shown as well as some IoT-related function blocks. In this example, data acquisition part and signal processing part are implemented in IoT Gateway (GW) in order to be closer to acoustic sensors (microphone array). Analytics part is organized into streaming analytics function box in IoT server. Here, data sources can be USB type microphone, analog microphone + data acquisition system (DAQ), or file format (WAV). In case of USB microphone (array), PyAudio [9] supports many kinds of configuration for data acquisition as python libraries. These days, many DAQs are commercially available which supports and provides python libraries (e.g. [10-11]), which are very useful when it comes to combination of OSS technologies provided by python codes especially in data science area (e.g. [12]). Acquired data would be transferred to preferable function boxes in signal processing part according to pre-communicated and pre-defined configuration. In this example, MVDR beamformer and some band-pass filters are selected, which show preferred behaviors as indicated by configuration scheme. gRPC server and client are implemented in IoT GW and server respectively. gRPC server is in charge of not only communication interface with gRPC client, but also actual distribution of configuration information from gRPC clients to each function block. For example, configurable information includes data source, sampling frequency, chunk size for streaming data, bit width, encoding, or more in data acquisition part. On the other hand, filter type and related parameters / coefficients, beamformer type, or more can also be configured. In this way, gRPC client can control data acquisition part and signal processing part. Furthermore, the order of processing can also be configured. As a supplement, Fig. 4 describes post data processing after some analytics function modules, where data would be transferred to some message broker, data bases, and applications (e.g. dashboard) with considering IoT-related consideration such as load balancing.

3. Verification

In order to verify scalability of the developed acoustic streaming data processing system for multiple applications, in this chapter verification results of enabling multiple applications (analytics) are shown with two example applications of machine state classification and anomaly detection of a motor. A MEMS-type 8-channels circular synchronized microphone array with 16 kHz sampling rate and int 16 sound encoding was utilized for both use cases [13] (Fig. 5).

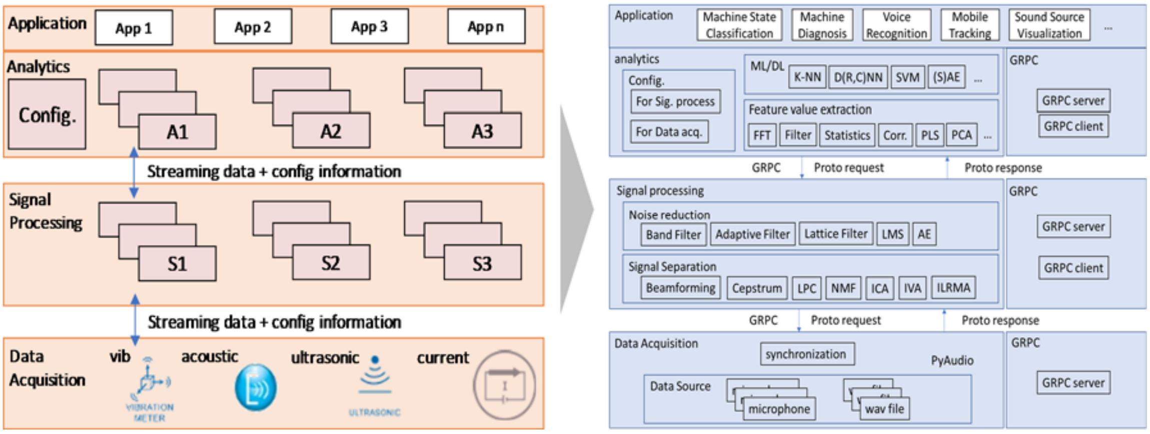


Fig. 2. Proposed system architecture (left), and its implementation on acoustic sensing system (right).

Property	Contents	Description	Direction
Command	STOP_STREAMING	STOP_STREAMING	Req.
	START_STREAMING	START_STREAMING	Req.
	SUSPEND_STREAMING	SUSPEND_STREAMING	Req.
	RESUME_STREAMING	RESUME_STREAMING	Req.
Config.	FilterType	Processes in signal processing part	Req./Resp.
	SamplingFreq	Sampling frequency	Req./Resp.
	ChunkSize	Streaming data size	Req./Resp.
	BitWidth	Number of bit per sample	Req./Resp.
	Encoding	Data type	Req./Resp.
Stream	AudioStream	Streaming data with requested property	Resp.

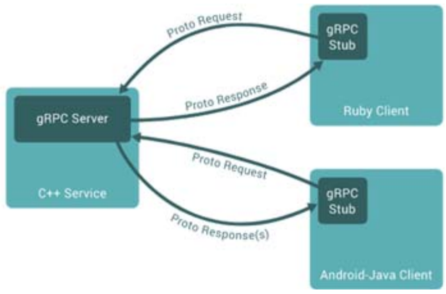


Fig. 3. Request/Response property on protocol buffer, and gRPC procedure [3].

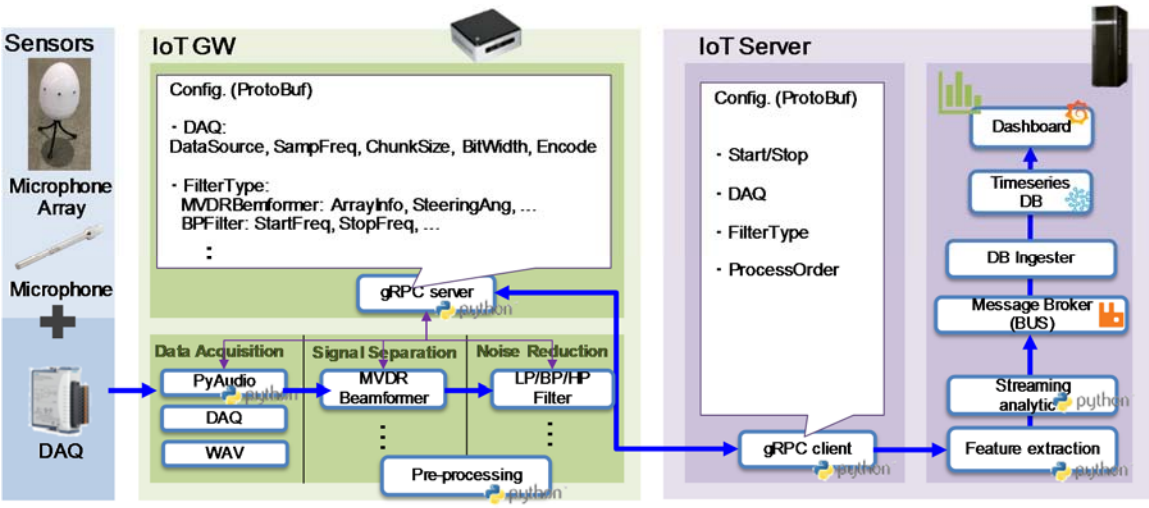


Fig. 4. A simple example implementation of proposed acoustic sensing system.

Fig. 5. Motor (left) and microphone array [13] (right) for verification.

The objective here was to verify seamless operationalization of this architecture with multiple applications. We successfully deployed relevant function modules for multiple applications on the same one sensing system to enable each application with one-stop configuration manner as shown in Fig. 6, which snows acoustic sensing system architecture with modules used in machine state classification and anomaly detection.

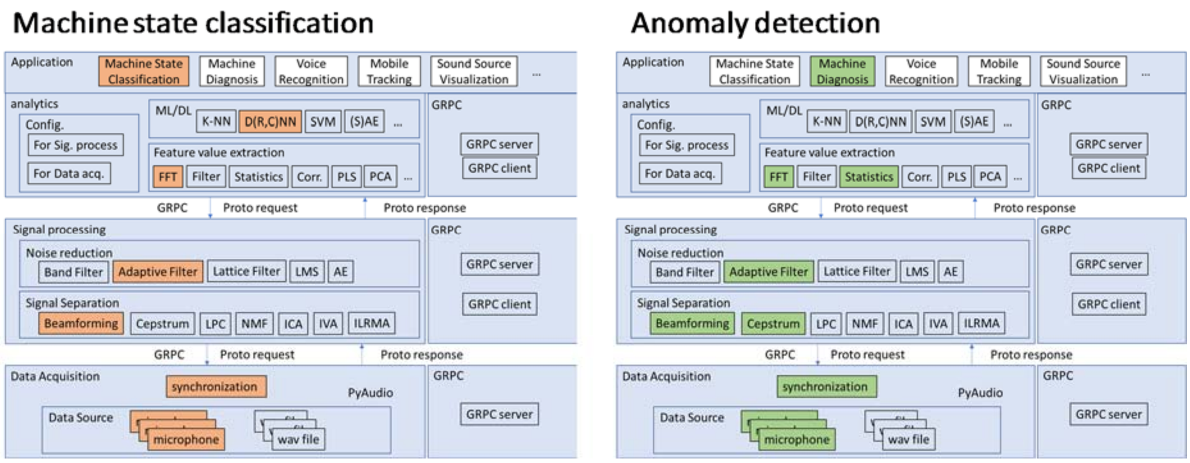


Fig. 6. Acoustic sensing system architecture with modules used in machine state classification and anomaly detection.

3.1. Machine State Classification

To verify the capability of enabling this application, we used an industrial motor for manufacturing field site. We used 'stop' state, 'low' speed state, and 'high speed' state as 3 different states of the motor. Fig. 7 shows spectrogram and averaged FFT results from the acoustic signal. A little difference can be seen between states.

With taking the spectrogram and averaged FFT into account, we developed a 4 layers neural network which consists of an input layer, two hidden layers, and an output layer, whose number of nodes are 129, 128, 64 and 3 respectively, where sequential model was utilized. The second and third models were based on sigmoid function, and final one on softmax function. This neural network was simple all connected neural network developed with OSS Keras / TensorFlow [14-15]. In this ML process, only the result of STFT was utilized as input of neural network, and out output nodes correspond to each state of the motor (stop, low, high).

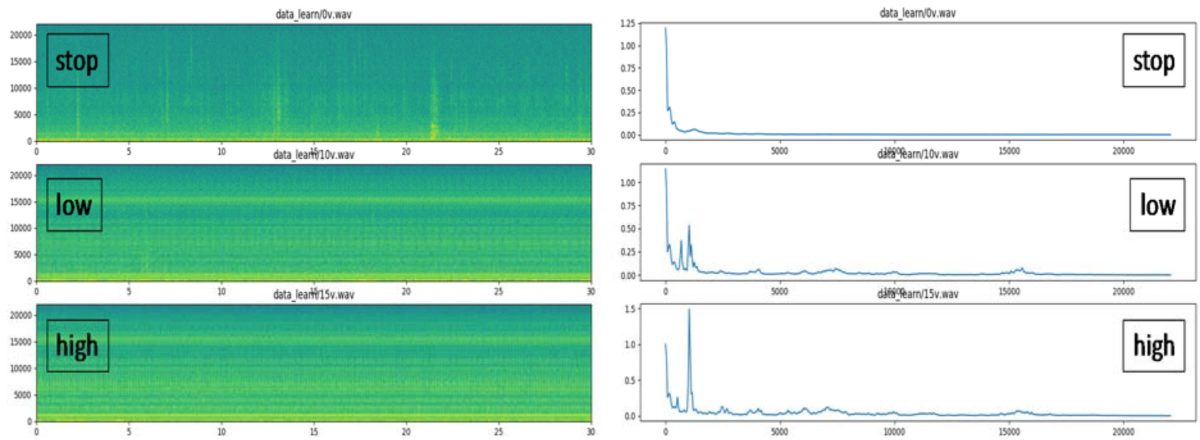


Fig. 7. Spectrogram (left) and averaged FFT results (right).

The classification results of streaming signal corresponding to each motor state are shown in Fig. 8, which shows 100 % classification accuracy. Classification was executed, then the result was output every 1 second. Streaming chunk size was 1024, which is actually the same size as STFT.

3.2. Anomaly Detection

The difference between "normal" and "anomaly" was demonstrated by the extent of tightening a screw which makes the motor mounted on a metal base plate, where anomaly state was defined as the motor with loosed screw. Fig. 9 shows raw waveform data of normal and anomaly status, where no significant difference can be found. Fig. 10 shows spectrogram and averaged FFT of normal and anomaly signals derived from a microphone respectively. Some differences between normal and anomaly around 200 Hz and 2,000 Hz can be found. Here, acquired synchronized raw signals were processed on a MVDR beamformer to make the beam focus on the motor.

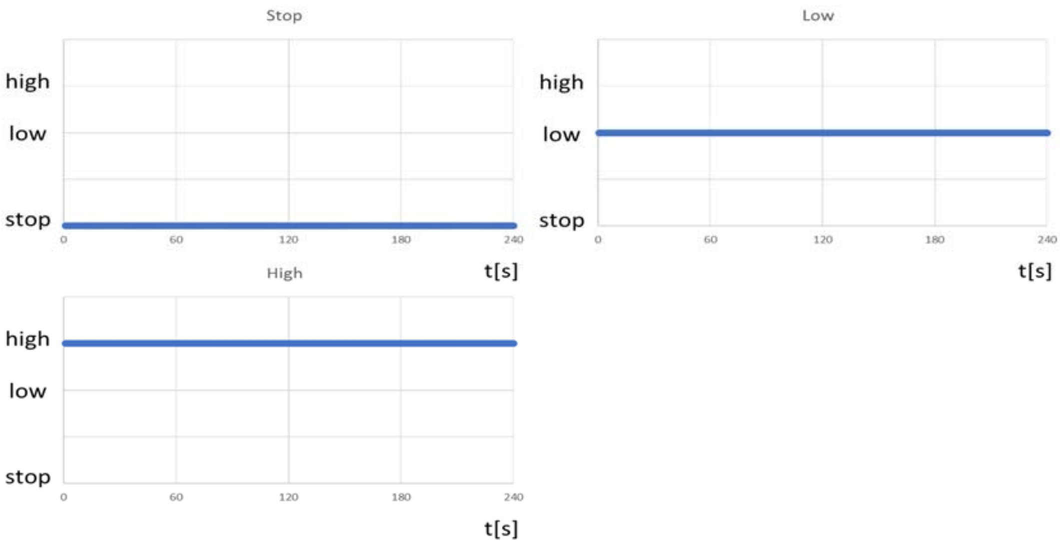


Fig. 8. Classification results.

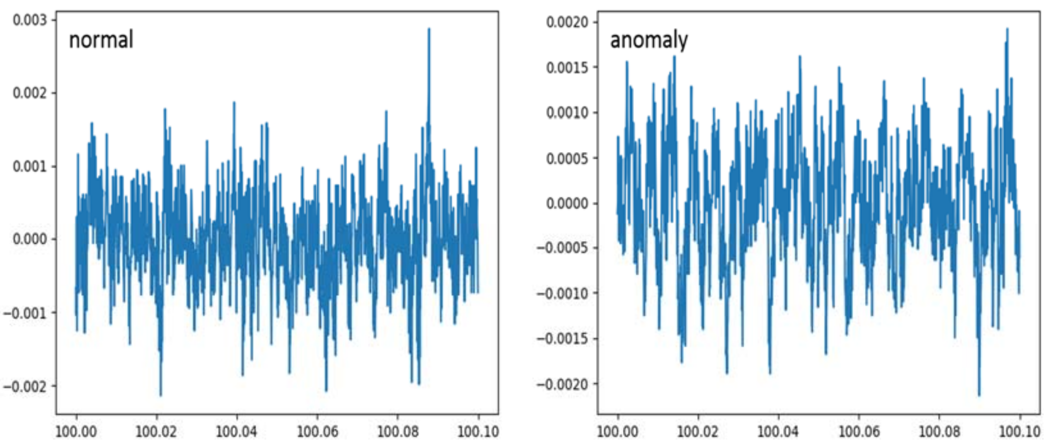


Fig. 9. Raw data of normal/anomaly signal.

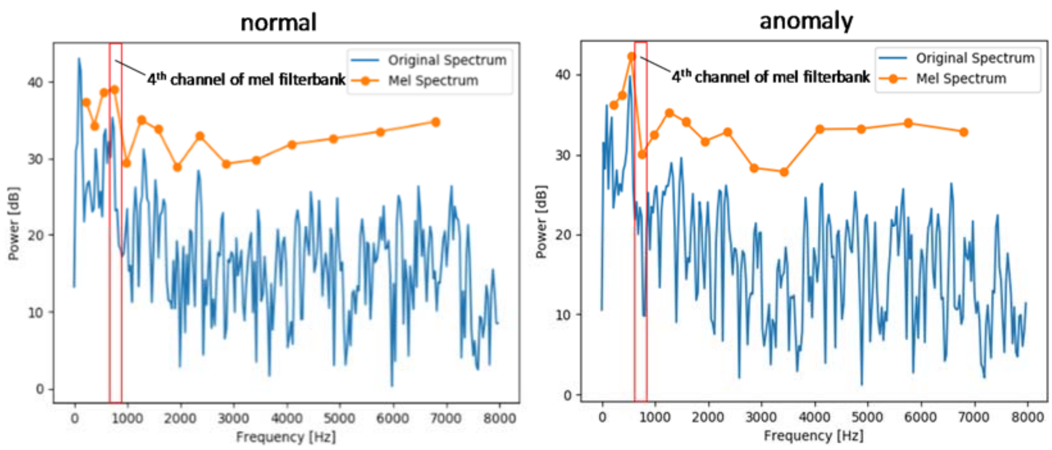


Fig. 10. Spectrogram and averaged FFT of normal and anomaly signals.

Then, the supervised NMF (nonnegative matrix factorization) with multiple bases, was applied to extract the clearer motor sound signal separated from other noise sources. Finally, in this application, GMM (Gaussian Mixture Model) was applied for training in order to judge anomaly degree (score) of normal and anomaly signals. In GMM process, mel filterbank was applied on the power spectrum from FFT results in order to have clearer difference between normal and anomaly. Regarding to feature values, which are independent variable input to GMM, power spectrum with 15-channels mel filterband and time-series difference (delta feature value) of power spectrum were applied. Therefore, a feature value vector with 30 elements was utilized to extract anomaly degree of each signal. Here, although many sound signal separation and data clustering technologies have been reported [16-20], relatively simple methods were applied in order to make system verification simple and include basic data processing components into streaming data processing system.

In Fig. 10, there seems to be some difference between normal and anomaly state on the 4th channel of mel filterbank. However, as shown in Fig. 11 and Fig. 12, distribution of power spectrum and delta value of 4th mel channel actually shows no clear separation of normal and anomaly signals. Fig. 11 shows distribution of normal signal with fitting result of one Gaussian distribution, and in Fig. 12 distribution of normal and anomaly signals with fitting result of two Gaussian distribution is shown.

Fig. 13 shows anomaly degree of original overall waveform (600 sec) of normal and anomaly as a result of this sequence of signal processing. This result shows obvious higher anomaly degree in case of anomaly state than normal state, which indicates that anomaly defined in this verification test can be successfully detected with this method. Additionally, this result can be obtained in time series manner (not static manner), which means developed system can provide real-time feedback of anomaly detection to manufacturing field site.

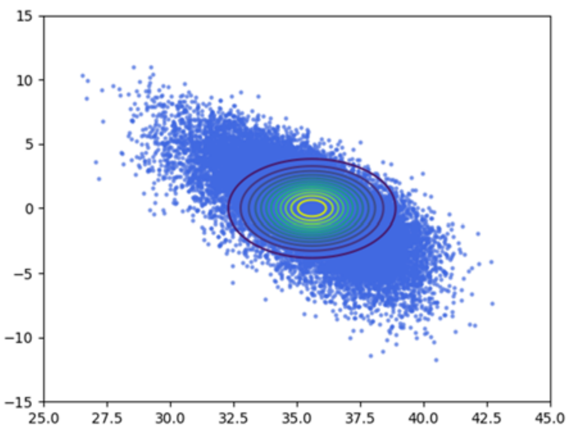


Fig. 11. Distribution of normal signal with fitting result of one Gaussian distribution. (Horizontal: power [dB], Vertical: delta [dB])

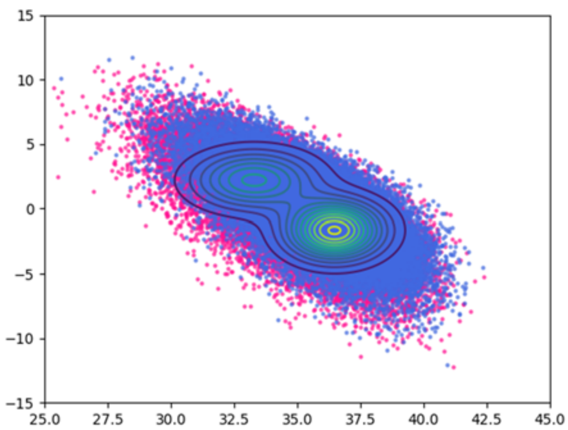


Fig. 12. Distribution of normal and anomaly signals with fitting result of two Gaussian distribution. (Horizontal: power [dB], Vertical: delta [dB])

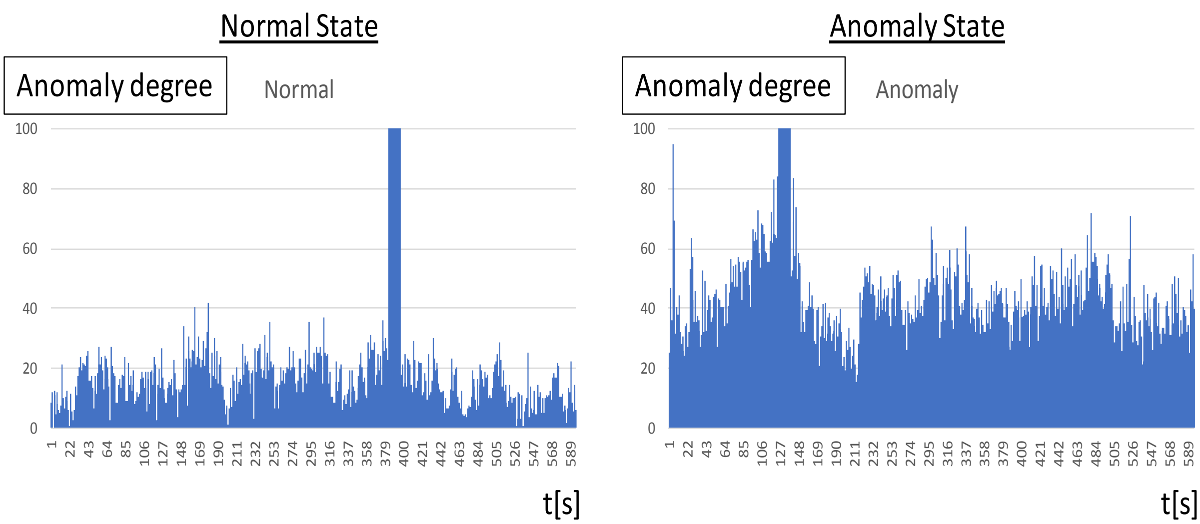


Fig. 13. Anomaly degree of overall waveform (600 sec) of normal and anomaly signal.

4. Conclusions

According to current challenges in manufacturing field site, especially unscheduled downtime and lack of skilled workers, we have proposed and designed a loosely coupled sensing system architecture, which is followed by some trial implementations into industrial IoT field. Acoustic sensors were applied as the first data source for scalable and reusable sensing system, which supports streaming data processing capability. The developed system has pluggable system architecture where data acquisition part, signal processing part, and analytics part can be mutually communicated with common APIs, which provides flexible combination of streaming data processing in multiple parts. The scalability of developed system was demonstrated on two example applications of machine state classification and anomaly detection. The successful operation of both applications on the same acoustic sensing system indicates that the loosely coupled / pluggable system architecture works as expected, which suggests the developed sensing system can also reduce time and cost for system integration related to applications establishment significantly.

References

[1]. Salonen A., Tabikh M., Downtime Costing—Attitudes in Swedish Manufacturing Industry, in *Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM'15),* 2016, pp. 539-544.

[2]. Radosław Wolniak, Downtime in the Automotive Industry Production Process – Cause Analysis, *Quality Innovation Prosperity (Kvalita Inovacia Prosperita),* Vol. 23, Issue 2, 2019.

[3]. Gu K., Stoyanov A., Skills, population aging, and the pattern of international trade, *Review of International Economics*, Vol. 27, Issue 2, 2019, pp. 499–519. https://doi.org/10.1111/roie.12386, Wiley Online Library.

[4]. Skevi A., Szigeti H., Perini S., Oliveira M., Taisch M., Kiritsis D., Current Skills Gap in Manufacturing: Towards a New Skills Framework for Factories of the Future, in Grabot B., Vallespir B., Gomes S., Bouras A., Kiritsis D. (eds.), Advances in Production Management Systems. Innovative and Knowledge-Based Production Management in a Global-Local World, APMS 2014, IFIP Advances in Information and Communication Technology, *Springer*, Berlin, Heidelberg, Vol. 438, 2014.

[5]. S. Pasha, *et al*., A Deep Learning Approach to the Acoustic Condition Monitoring of a Sintering Plant, in *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit,* 2018.

[6]. E. Wangkanklang, *et al*., The Progress Investigation of Ishikawa Mixing and Grinding Machine Using Sound Signal Processing, in *Proceedings of the IEEE 7th Global Conference on Consumer Electronics (GCCE)*, Nara, Japan, 9-12 Oct. 2018.

[7]. Y. Serizawa, Y. Shomura, A Flexible Acoustic Sensing System and its Application to IIoT – Manufacturing Field Site, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'2019)*, Tenerife (Canary Islands), Spain, 25-27 September 2019, pp. 18-23.

[8]. gRPC/Proto buffer, Google, https://grpc.io/

[9]. PyAudio, https://pypi.org/project/PyAudio/

[10]. National Instruments, NI-DAQmx, https://nidaqmx-python.readthedocs.io/en/latest/

[11]. LabJack, LabJack T7, https://labjack.com/support/t7-software-options

[12]. Giannakopoulos T., pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis, *PLoS ONE*, Vol. 10, Issue 12, 2015. e0144610. https://doi.org/10.1371/journal.pone.0144610

[13]. TAMAGO-03, System in Frontier Inc., http://www.sifi.co.jp/system/modules/pico/index.php?content_id=39

[14]. Keras, https://keras.io/

[15]. TensorFlow, https://www.tensorflow.org/

[16]. Saeed Aaqib, Urban Sound Classification, Part 1., Aaqib Saeed. September 3, 2016. Accessed November 25, 2017. http://aqibsaeed.github.io/2016-09-03-urban-sound-classification-part-1/.

[17]. Shaikh Faizan, Getting Started with Audio Data Analysis (Voice) using Deep Learning, *Analytics Vidhya*, 23 August 2017. Accessed November 25, 2017. https://www.analyticsvidhya.com/blog/2017/08/audio-voice-processing-deep-learning/.

[18]. M. Togami, *et al*., Noise robust speech dereverberation with Kalman smoother, in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,* Vancouver, BC, Canada, 26-31 May 2013.

[19]. R. Takashima, *et al*., Separation of vibration-derived sound signals based on fusion processing of vibration sensors and microphones, in *Proceedings of the 25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 28 Aug.-2 Sept. 2017.

[20]. Y. Kawaguchi, *et al*., Non-Negative Novelty Extraction: A New Non-Negativity Constraint for NMF, in *Proceedings of the 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, Tokyo, Japan, 17-20 Sept. 2018.

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 57-63

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Methodology for the Choice of a Right Digital Magnetometer for the Vehicle Detection Purposes

[1, *] Michal HODOŇ, [1] Ondrej KARPIŠ, [1] Peter ŠEVČÍK, [1] Juraj MIČEK, [1] Peter ŠARAFÍN, [1] Martin HÚDIK and [2] Andrea KOCIÁNOVÁ

[1] University of Žilina, Department of Technical Cybernetics, Univerzitná 8215/1, 01026, Slovakia
[2] University of Žilina, Department of Highway Engineering, Univerzitná 8215/1, 01026, Slovakia
[1] Tel.: +421415134355
[*] E-mail: michal.hodon@fri.uniza.sk

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: The paper is aimed at the problematic of vehicle detection, which is based on the employment of standard digital anisotropic magneto-resistive magnetometer sensors. There is a number of suitable sensor modules available in the market, which can be used for this purpose. The sensors, however, differ in parameters, such power consumption, maximal sampling frequency, measurement range, or measurement sensitivity. The aim of this paper is to provide a kind of coherent comparison of such as sensor performance. The bank of sensors is tested first in the laboratory environment and then, the sensors with the best results are further tested in the outside scenario. During the test, the needs of the final application were stressed, that the proper sensor module will be chosen. This paper is an extended version of the conference paper [1], with an added section dedicated to the outside measurements.

Keywords: Magnetometer, Sampling, Performance, Power Consumption, Vehicle Detection.

1. Introduction

The world of information technology is now more than ever characterized by the interconnecting of individual systems. According to IoT Analytics [2], more than 17 billion devices are currently connected, of which about 7 billion are the Internet of Things (IoT). It should be noted that smartphones, tablets, and laptops are not included in this amount. The number of active IoT devices is expected to rise to 10 billion by 2020 and t22 billion by 2025.

The largest number of devices are currently connected via short-range technologies (e.g. Bluetooth, ZigBee, Z-wave).

In summary, these technologies are referred to as Wireless Personal Networks (WPAN). It is expected that a large part of the future growth of IoT devices will be provided by the development of low-power Wide Area Networks (LPWANs). The main standards in this area (Sigfox, Lora, NB-IoT) currently account for more than 25 million connected devices, up to 2 billion by 2025. Very low power consumption (i.e. long battery life, possibility to use green power supply) and a very large range that can be more than 20 km promises to increase to a large number of applications. All these technologies use mainly free frequency bands for communication. The only way to achieve growth with limited resources such as the frequency band is to use it as efficiently as possible [3].

Most of the IoT data goes to public or private cloud storage where are saved, processed and visualized. This trend is unlikely to change in the future. But what must be changed is the qualitative content of data transmitted to cloud servers. Due to the limited capacity of wireless technologies, it is impossible to transfer all data in a raw, unprocessed state. The further, the greater it is the importance to have the data preprocessing directly at the capture place. Moving data processing closer to its place of origin (of moving applications and functionalities closer to the user) is currently a much-discussed topic and is generally referred to as “Edge computing”. The importance of this trend in the IoT field is also evident from the activities that the largest cloud companies such as Microsoft, Google and Amazon are developing in this area. As an example, Google Edge TPU is a processor that is focused on using machine learning and artificial intelligence “on the edge” [4].

Thanks to the continuous development of technologies in all areas of electronics (from sensors, through microcontrollers to radio modules), the number of IoT applications is growing rapidly. Lower power consumption, as well as higher computing power for even the simplest processors, also allows the use of sensor types that could not be used in the past due to the large amount of data they produce. This applies mainly to image and sound recording but also to other signals with higher data volume (vibration measurement, magnetic field changes, etc.).

The relatively high computational power available at the place of signal acquisition allows the data volume to be reduced to an acceptable level with respect to the communication channel used.

2. Problem Description

Vehicle detection technologies were massively developed in the last couple of decades. Their developments went hand in hand with the requirements, which are currently put on the measurement strategies, where the accuracy is viewed as a necessity. According to the resources available, different measurement methods can be applied, with different results reached. For example, with camera and image recognition techniques, different statistics can be derived, such as the number of vehicles of different types, speed averages, problematic drivers, etc. Problems only occur, where the measurement scenario provides only a limited number of resources, but the measurement requirements are still high. For this purpose, it is necessary to build special sensory devices.

Because most of all vehicles have noteworthy measures of steel material in their body, AMR sensor magnetometers represent a good technical mean for identifying vehicles. As mentioned, authors in [5], the application of such sensors for the purposes of traffic monitoring provides a cost-effective and easy-to-install solution. After the application of Discrete Fourier Transform and Principal Component Analysis (PCA) for extracting features from magnetic signatures, the detected vehicles can be easily classified into three different categories. Beside not so demanding application as is the detection of the static vehicle at, for example, parking lot [6], the most applications ask for the real sensor performance, so just the features with time-stamp can be extracted from the measured magnetic field. It is therefore evident, that signal pre-processing at the sensor side, or “sensor edge”, is very important.

Different approaches can be applied at the sensor side, with different performance requirements. For example, authors in [7] proposed and implemented a simple algorithm based on a nonlinear transformation of the measurements as correlation-based driving direction classification. In [8], a simple five-state machine process algorithm was developed and implemented for real-time vehicle detection and counting. Speed estimation there relied on precise time-stamping of vehicle arrival and departure instants, which depended on T-sync algorithm accuracy. However, sometimes it is necessary to add some other supporting sensors to increase the measurement accuracy. In this case, particle filtering can play the role. In [9], an accelerometer sensor was added into the system. The paper showed that this approach can serve as a kind of feasible solution. For the sensor fusion, multi-rate particle filter was used.

Examples above proved that AMR sensors can be considered as reliable, low-power and cost-effective technology, which is suitable for vehicle detection. An integration of AMR sensors into the traffic infrastructure opens the space for novel applications, aimed at the topic of intelligent transportation systems (ITS). These sensors can be used for example for the traffic monitoring (traffic volume, gap times, speed) or intelligent traffic control including the real-time adaptive control of the traffic lights [10].

3. Structure of Magnetic Traffic Counter

As was mentioned within the examples above, the utilization of AMR sensors would be very tough without signal preprocessing at the sensor side, or eventually, signal post-processing at the side of the final application - magnetic traffic counter. Therefore, though the AMR sensor can be considered as a very important part of the final application, there are also other parts (Fig. 1):

- Control unit – this is a crucial part of any ITS sensor because it controls all algorithmizing processes. Parameters such as floating-point operations per second (FLOPS), as well as power requirements, are the key features to be considered when selecting an appropriate controller.

- Power – in most application scenarios, the power cords are missing, therefore requirements on power efficiency are very important. Different energy harvesting techniques are obviously used as supporting technologies for the sensor lifetime extension.

- Memory – in applications where the data are stored at the sensor side, it is necessary to integrate power-efficient solution for the data storage (SD card 50 mA @ 3.3 V, SPI Flash (25 mA @ 3.3 V).

- Communication subsystem – communication with user or with parent application should be primarily used for the data download, secondary for the settings adjustments. Communication can be event-based (button, shake,) or synchronous waking and sending the data in a defined time window. Wireless (Bluetooth, Xbee, ISM proprietary protocols...) or wired (USB, RS485, Ethernet...) means can be used for these purposes.

- Sensor subsystem – beside magnetometers, which are going to be described in another section, other supporting sensors could take place according to the application specifics. For example, GNSS sensors for timing purposes, sound sensors for the classification of voices, or gas sensors for the measurements of environmental conditions.

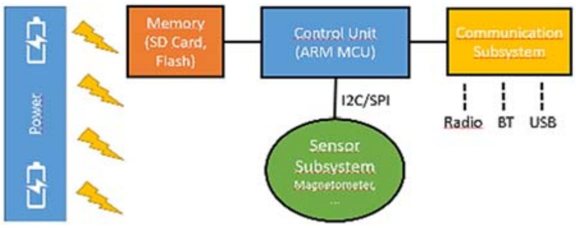


Fig. 1. ITS sensor structure.

4. Magnetometer Selection

There are several digital AMR magnetometers as SoC applicable for ITS applications available at the market with different performance and quality measures. To select the right one, first, the application scenario must be described.

The desired application scenario comes out from the project "Innovative monitoring and analysis of traffic on the cross-border road network", which is aimed at the development of special low-cost, low-power, wireless, portable magnetic traffic counter for the analysis of the traffic volume at the cross-border area between the Slovakia and the Czech Republic. This area can be described with mountains and forests, where the electricity is provided rarely. The proper sensor should be able to detect the disturbances of the magnetic field relied on its proximity, and thus to detect the vehicle size and direction (Fig. 2).

Fig. 2. Earth's magnetic field through vehicle [20].

For this purpose, a market analysis was performed, resulting in the choice of 10 magnetometers for further evaluation. The list of chosen magnetometers is provided in the Table 1, together with the basic parameters which were extracted from [11-20]. We were focused on chips, which are without any problems available at the market either from re-sellers or from specified suppliers in a number of single pieces.

Table 1. List of the tested magnetometers [11-20].

	Magnetometer Type	Output Noise	Sensitivity [uT/LSB]	Measurement Range +- [uT]	Output Data Rate [Hz]	Current Consumption [uA]	Idle [uA]
1	**LSM303AGR**	0.3 uT RMS	0.15	5000	100/150	950/250 @ 100 Hz	2
2	**LIS2MDL**	0.3 uT RMS	0.15	5000	100/150	950/250 @ 100 Hz	1.5
3	**LSM303AH**	0.3 uT RMS	0.15	5000	100/150	950/250 @ 100 Hz	2.5
4	**LSM303C**	0.35 uT RMS	0.058	1600	80	270/40 @ 20 Hz	6
5	**LSM9DS1**	0.35 uT RMS	0.014 - 0.058	1600	80	600 @ 20 Hz	6
6	**MPU9250**	N/A	0.15	4900	100	280 @ 8 Hz	3
7	**FXOS8700CQ**	<100 nT/Hz @ 100 Hz	0.1	1200	800	575 @ 400 Hz	2
8	**MLX90393**	varies acc. to settings	0.16 - 5.8	4800	700	100 @ 10 Hz	2.4
9	**RM3100**	30-15 nT	0.05 - 0.013	800	1600/n	70-260 @ 8 Hz	1
10	**HMC 5983**	N/A	0.07 - 0.40	100-800	220	100 @ 7.5 Hz	2

From the reference sources is also obvious, that we were focused on the magnetometers with digital output for minimization of power consumption. From the table is evident, that there are some parameters, which can be left out from the comparison since they are almost the same among all of the sensors. What is important in terms of the aimed measurements is the following:

- Output noise – lowest (around 0.3 uT and less) – to enable the sensor to measure accurately, with low own noise, and invulnerable to any noise source,

- Measurement sensitivity – highest (around 0.1 uT/LSB and higher) – to enable the sensor to sense any change of magnetic field properly,

- Measurement range – low – to provide the best measurement quality to sense with the highest sensitivity in relatively low measurement range,

- Output data rate – highest – to sample the data at the highest possible speed (not to lose any information),

- Power consumption – lowest – power consumption of the whole device is very important.

What is however also important, and what could be easily investigated in laboratory conditions is the measurement repeatability, resp. the spread of measurement data of each magnetometer axis under the same laboratory conditions. In terms of measuring the variability of data spread, the standard deviation S is the preferred and most used measure as a square root of the variance (1):

$$S_K = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}|M_{Ki} - \mu_K|^2}, \qquad (1)$$

where μ_K is the mean of the measured value in all three axes X, Y, Z of magnetometer of N measures, where $K \in \{X, Y, Z\}$ (2):

$$\mu_K = \frac{1}{N}\sum_{i=1}^{N} M_{Ki} \qquad (2)$$

Interpretation of the coefficient of variation (CV) for each magnetometer can be then used to interpret the variability of a sample relative to its mean by dividing the standard deviation by the mean (3):

$$CV = \frac{S}{|\mu_K|} \qquad (3)$$

To properly measure the data, as to interpret the inputs for (1), (2) and (3), the special measurement set was developed and installed as can be seen in the block schematic in Fig. 3.

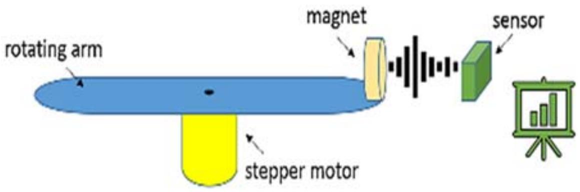


Fig. 3. Block schematic of the measurement set.

The measurement scenario consisted of the rotating arm which was fixed on the precisely driven stepper motor which was controlled through the AtMega8 MCU and two H-bridges. On the arm was placed strong neodymium magnet which was moved across the defined track. The track consisted of 42 measurement points, on which the magnet stops for exactly 1200 ms. Every tested magnetometer from the group available in Table 1, was place exactly on the same position fixed at a certain distance of 1 cm from the magnet installed at the rotating arm. So, during the measurement of each magnetometer, the same measuring conditions were achieved:

- The same distance between the moving magnet and the statically placed measuring unit,
- The same movement of rotating arm, characterized with the same speed and the same track,
- The same position of magnetometer relative to the rotating arm.

The magnetometers were set to the maximum output data rates with no filter implemented, thus the real raw data could be measured. All magnetometers were interconnected through the I2C interface with Raspberry Pi v 3.0, through which all settings were adjusted and where all data were recorded. The influence of rotating magnet to the magnetometer can be seen in Fig. 4.

The whole test scenario in real can be seen in Fig. 5.

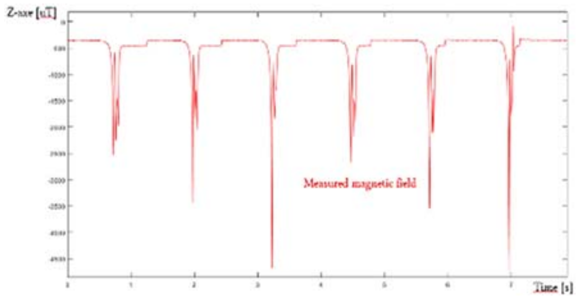


Fig. 4. Influence of rotating magnet to the magnetometer.

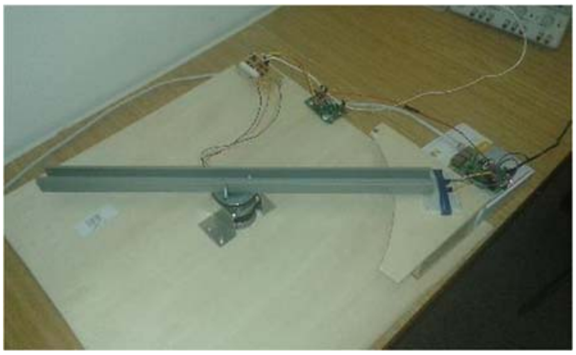


Fig. 5. Test scenario.

The measurement results can be seen in Table 2 as the interpretation of the coefficient of variation for X, Y and Z axes of every tested magnetometer.

Table 2. Measured results.

	CV(X)	CV(Y)	CV(Z)
LSM303AGR	0.0750	0.0075	0.0332
LIS2MDL	0.0748	0.0073	0.0331
LSM303AH	0.0748	0.0073	0.0332
LSM303C	0.0209	0.0037	0.0100
LSM9DS1	0.0207	0.0038	0.0149
MPU9250	0.0237	0.0062	0.1391
FXOS8700CQ	0.1010	0.0441	0.0221
MLX90393	0.2560	0.2837	0.0583
HMC 5983	0.0273	0.0086	0.0283
RM3100	0.0031	0.0042	0.0106

From the results is evident, that magnetometers LSM303AGR, LIS2MDL and LSM303AH have probably the same AMR core. From the results is also evident that LSM303C, LSM9DS1, FXOS8700CQ and RM3100 provide the best results. From the selected modules above, two sensors were chosen for further investigation in the scenario of real traffic - FXOS8700CQ and RM3100. Though RM3100 is quite costly, and has the biggest size, it provided really the best results. This is because it is based on the different technology as opposite to AMR – proprietary magneto-inductive technology. However, its output data rate depends on the settings – its sensitivity depends on the number of sensed channels. FXOS8700CQ was preferred over the LSM family mainly because of its high output data rate (800 Hz vs. 200 Hz), even it provided worse measurement results (Table 2).

5. Testing in the Scenario of Real Traffic

For the purpose of measurement, the special testing scenario was set up in the village Porúbka, which is situated on the road of I[st] class I/64 in Slovakia. The road connects nearby towns Žilina and Rajec with relatively high traffic load in terms of Slovakia. The measurement scenario consisted of the special construction, on which HD camera together with sensors and control unit was fasten. The block schematic of the measurement set-up can be seen in Fig. 6.

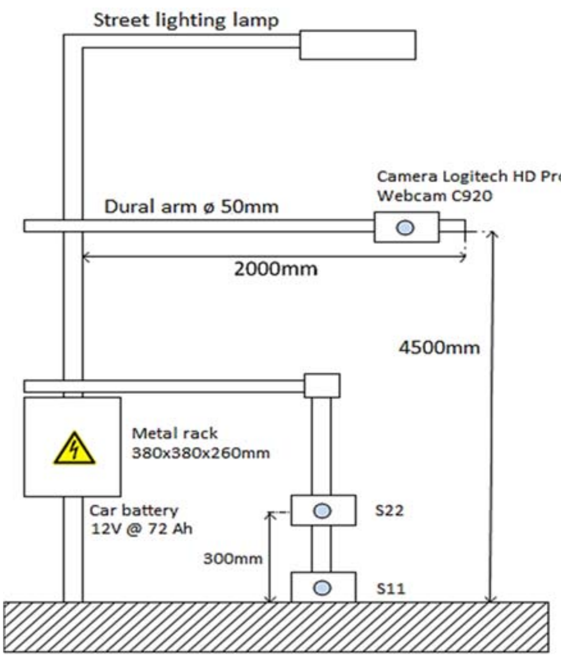


Fig. 6. Block schematic of the measurement set-up.

As it is obvious from the schematic, for the measurement purposes, two sensor modules – S11 an S22 were introduced. Both modules had the same structure – they consisted of one RM3100 sensor and one FXOS8700CQ, both interconnected with Raspberry Pi v 3.0 single-board computer through the I2C bus. The modules shared the same physical bus, which served enough throughput for the maximal sensor sampling. Raspberry Pi was used as the data logger. For better imagination, in Fig. 7 and Fig. 8, there is displayed the whole measurement scenario in detail.

Fig. 7. Measurement scenario in Porúbka.

Fig. 8. Measurement scenario in Porúbka.

From the figures is evident, that sensor modules were installed in two levels. The first sensor module, named S11, was installed in parallel with the pavement ground, the second sensor module was installed at a height of 30 cm. As it was also mentioned, all sensors were set to maximal output data rates with maximum possible sensitivity in all three axes of the magnetometer. Then the recorded raw data were analyzed with simple band-pass filter applied, as for example in [5].

To keep the measurement integrity, the camera recording of actual traffic was performed in parallel with the magnetometer measurement. Measured data could be then in post-processing compared with the extracted images of traffic flow. The results can be seen in Fig. 9. In the upper images, there are displayed figures of two vehicles extracted from the camera record – the van and the car which drove in line just behind. In the lower pictures can be seen the measured data. From the figures is evident, that RM3100 served much better results – as in measurement sensitivity, as in output noise.

In Fig. 10, the comparison of the sensor performance, when placed in the ground vs. 30 cm above the ground was done on the example of truck, car and van sequence. From the figure is evident, that slightly better results are provided, when the sensor is placed a bit above the pavement ground when the sensor is placed aside the road.

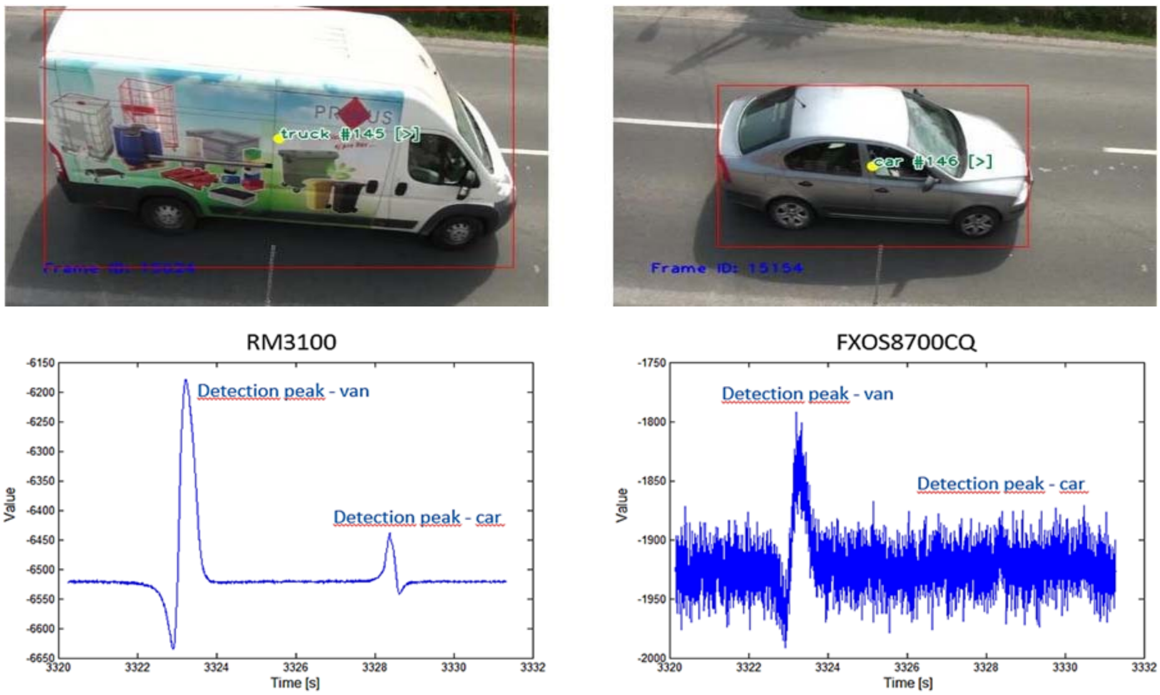


Fig. 9. Measured data – comparing the performance between RM3100 and FXS8700CQ in the S11 scenario.

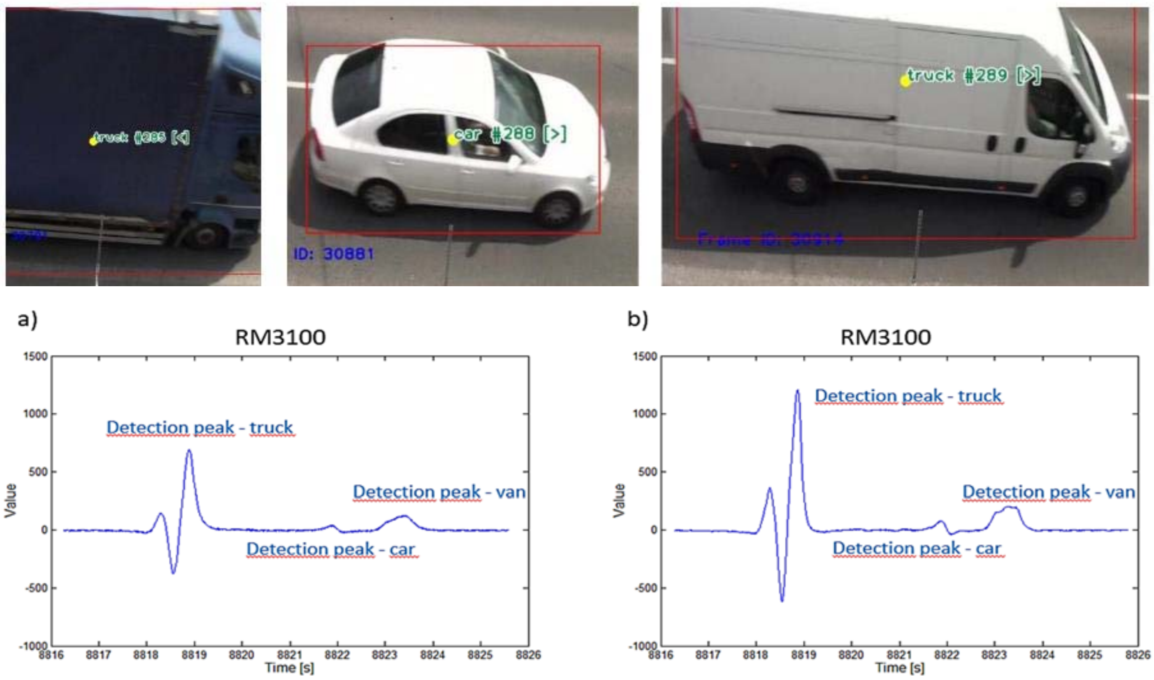


Fig. 10. Measured data – comparing the performance of RM3100 between a) S11 and b) S22 scenario.

6. Conclusion

In this paper, the basic selection and performance investigation of market-available magnetometers was made. The sensors were selected according to the final application requirements which are summarized below:

- Non-invasive;
- Reliable;
- Long working time;
- Low power consumption;
- Vehicle presence;
- Vehicle type;
- Easy & fast set up.

According to measurement results in the laboratory, the smaller groups of magnetometers for further evaluation were selected. In the selection, we considered output data rates of single magnetometer together with the measurement repeatability as the important parameters, which serve the additional information about the quality of the measured signal. After the laboratory tests, we continued with the testing in the scenario of real traffic. From the measurement results, we concluded that the best sensor which most suited the needs of the final magnetic traffic counter application is RM3100. In future research, the first prototype built on the selected sensor will be introduced and further tested in the application scenario.

Acknowledgment

This contribution is the result of the project Innovative monitoring and analysis of traffic on the cross-border road network, ITMS 304011P777 supported by the Interreg V-A SK-CZ Cross-border Cooperation Program.

INTERREG V-A SLOVENSKÁ REPUBLIKA ČESKÁ REPUBLIKA SK-CZ 14-20

EURÓPSKA ÚNIA EURÓPSKY FOND REGIONÁLNEHO ROZVOJA SPOLOČNE BEZ HRANÍC

References

[1]. M. Hodoň, *et al.*, An Overview of Magnetometer Sensors Performance for the Purposes of Traffic Flow Monitoring. in *Proceedings of the 5th International Conference on Sensors and Electronic Instrumentation Advances*, Tenerife (Canary Islands) Spain, 25-27 September 2019, pp. 195-199.

[2]. State of the IoT 2018: Number of IoT devices now at 7B – Market accelerating. https://iot-analytics.com/state-of-the-iot-update-q1-q2-2018-number-of-iot-devices-now-7b/

[3]. K. Mekki, *et al.*, A comparative study of LPWAN Technologies for large-scale IoT deployment, ICT Express, Vol. 5, Issue 1, 2019, pp. 1-7. https://doi.org/10.1016/j.icte.2017.12.005.

[4]. Edge TPU, https://cloud.google.com/edge-tpu/ Sensors Web Portal (http://www.sensorsportal.com).

[5]. H. Tafish, W. Balid, H. H. Refai, Cost effective Vehicle Classification using a single wireless magnetometer, in *Proceedings of the IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, 5-9 Sept. 2016, Paphos, Cyprus.

[6]. A. Gadžović, N. Lekić, Igor Radusinović, Automation of parking services using UHF RFID technology and magnetometers, in *Proceedings of the 21st Telecommunications Forum Telfor*, Belgrade, Serbia, 26-28 Nov. 2013.

[7]. N. Wahlström, R. Hostettler, F. Gustafsson, W. Birk, Classification of Driving Direction in Traffic Surveillance Using Magnetometers, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, Issue 4, Aug. 2014, pp. 1405-1418.

[8]. W. Balid, H. Tafish, H. H. Refai, Intelligent Vehicle Counting and Classification Sensor for Real-Time Traffic Surveillance, *IEEE Transactions on Intelligent Transportation Systems*, Vol. 19, Issue 6, June 2018.

[9]. R. Hostettler, Petar M. Djurić, Vehicle Tracking Based on Fusion of Magnetometer and Accelerometer Sensor Measurements with Particle Filtering, *IEEE Transactions on Vehicular Technology*, Vol. 64, Issue 11, Nov. 2015.

[10]. K. Zaatouri, Mohamed Hechmi Jeridi, T. Ezzedine, Adaptive Traffic Light Control System Based on WSN: Algorithm Optimization and Hardware Design, in *Proceedings of the 26th International Conference on Software, Telecommunications and Computer Networks*, Split, Croatia, 13-15 Sept. 2018.

[11]. LSM303AGR Datasheet, DocID027765 Rev 10, 2018.

[12]. LIS2MDL Datasheet, DocID030621 Rev 5, 2018.

[13]. LSM303AH Datasheet, DocID027766 Rev 7, 2018.

[14]. LSM303C Datasheet, DocID024975 Rev 2, 2014.

[15]. LSM9DS1 Datasheet, DocID025715 Rev 3, 2015.

[16]. MPU-9250 Product Specification Revision 1.1, Document Number: PS-MPU-9250A-01 Revision: 1.1 Release Date: 06/20/2016.

[17]. FXOS8700CQ Technical Datasheet, Rev. 8, 25 April 2017.

[18]. MLX90393 Triaxis® Magnetic Node Datasheet, Revision 003, 14 September, 2017.

[19]. RM3100 & RM2100 Sensor Suite User Manual, Doc 1017252 R07, PNI Sensor Corporation.

[20]. HMC5983 Datasheet, Form # 900425, Honeywell Application Note – AN218 Vehicle Detection Using AMR Sensors, *Honeywell International Inc.*, September 2011.

___

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 64-71

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

LoRa-based Wireless Sensor Networks for Urban Scenarios Using an Open-source Approach

[1,*] Stefano TONDINI, [1] Simone TRITINI, [1] Matteo AMATORI, [2] Silvia CROCE, [3] Stefano SEPPI and [1] Roberto MONSORNO
[1] Eurac Research, Center for Sensing Solutions, 13/A Volta St., 39100 Bozen, Italy
[2] Eurac Research, Institute for Renewable Energy, 13/A Volta St., 39100 Bozen, Italy
[3] NOI Techpark Südtirol-Alto Adige, Tech-Transfer Digital, 13/A Volta St., 39100 Bozen, Italy
[1] Tel.: + 39 0471 055266, fax: + 39 0471 055099
[*] E-mail: stefano.tondini@eurac.edu

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: This work brings together two wireless sensor networks achieved within two European projects aimed at accomplishing novel approaches to urban monitoring. The networks share the same end nodes but differ in the architecture and the communication protocol, namely raw-LoRa and LoRaWAN (868 MHz). By the one hand, the raw-LoRa option allows for the full control on exchanged data, as the network management is implemented at hardware level. By the other hand, the LoRaWAN option provides a standard way to data gathering and comes with enhanced features for devices/applications management. The architectures rely on open-source compliant components both in terms of software (i.e. network management and data storage) and hardware (sensor-nodes and gateways). All these aspects have made it possible to achieve two flexible and low-cost platforms for environmental monitoring applications, among others. Data collected both outdoor and indoor enable us to assess the reliability and the drawbacks of the wireless sensor networks reported.

Keywords: LoRa, LoRaWAN, Low-power wide-area network, Sensor node, Gateway, Timeseries, Environmental monitoring.

1. Introduction

The use of low-power wide-area network (LPWAN) for environmental monitoring is taking hold because of their high energy efficiency and the long coverage range achievable [1]. Despite the limited data throughput, what makes this technology particularly attractive is the low implementation cost. However, communication robustness is still critical. Although gap filling methods and data transmission reduction schemes are becoming more and more satisfactory to cope with data lacks within timeseries [2-3], communication interference for wireless sensor networks (WSNs) remains a critical point to be addressed, especially in urban areas. In this work, we go through this aspect by considering two purpose-built WSNs, where the first one implements only the physical layer raw-LoRa as a communication protocol, while the second one encompasses the LoRaWAN application layer, as well [4]. These communication protocols lay into the Industrial Scientific Medical (ISM) frequency band (Europe 868 MHz)

Experimental data have been collected within the frame of two research projects, which are described in the following section.

2. Urban Monitoring Scenarios

2.1. SINFONIA Project

The European project SINFONIA [5], "Smart Initiative of cities Fully cOmmitted to iNvest In Advanced large-scaled energy solutions", is a five-year initiative to deploy large-scale, integrated and scalable energy solutions in mid-sized European cities. Technically and technologically innovative actions have been implemented in two pilot cities, Bolzano (Italy) and Innsbruck (Austria), with the aim to transform them into smart cities and to propose solutions easily replicable and adoptable by other European cities. Within the framework of SINFONIA, building refurbishment interventions to achieve high-energy performance and to improve interior comfort have undertaken for six residential complexes in the south district of Bolzano. Among those, the apartment complex located in Via Brescia-Cagliari. Refurbishment interventions on the building included, among others: external insulation with white plaster, substitution of the balconies' parapets, renovation of the roof by converting the space into new apartments and by adding an additional floor, and installation of a new solar thermal plant on the roof and on the south multifunctional façade.

After the completion of the construction works, the area is subject to an in-situ monitoring campaign, which will last for one year and collect data relatively to all the main meteorological parameters. The objective is double; first, to verify the impact of the renovation interventions – as, for example, the modification of the façade color from brown to white - on the microclimatic conditions in the district. The second aim is to collect data useful to calibrate the environmental used to simulate the current conditions (i.e. solar irradiation, microclimate, and airflow) in the area and the potential impacts of future modifications of its surfaces (e.g. installation of a green façade, creation of water urban features, etc.) [6-7].

2.2. BEACON Südtirol - Alto Adige Project

The BEACON Südtirol - Alto Adige project [8] is an ERDF (European Regional Development Fund) funded initiative that aims to create a favorable environment for the development of new ideas, applications and products through the installation of a Beacon Network and the setting up of an IoT Pilot Area at the Technology park of Bolzano (NOI Techpark). Moreover, the project wants to foster the creation of a local community of IoT experts, companies and research institutions that cooperates toward digital innovation in many different sectors.

Within the BEACON project, a technology and market analysis regarding the IoT state of art has been carried out, which has led to the release of a whitepaper in 2019 [9]. This first result, together with the feedback collected from the IoT community, suggest that LoRaWAN is one of the most interesting emerging technologies for IoT applications. Therefore, an experimental open-access LoRaWAN network has been made available within IoT pilot area of the NOI Techpark, with a view to:

- Create a free and easy-access infrastructure for developers and companies, where they can develop and test new applications, sensors, etc.:
- Foster the cooperation between research and industry at the level of co-creation approach;
- Facilitate the startup of innovation projects and the setup of new IoT-based services and products.

Next step of the project is to empower the coverage at the NOI Techpark with other promising IoT communication protocols, namely Sigfox, Symphony Link, 5G, among others. By doing that, it will be possible to compare the performances of the different technologies on a specific application and then choose the most suitable one for the final implementation.

3. Network Architectures

3.1. raw-LoRa

Within the SINFONIA project a 20 sensor nodes WSN has been set up and deployed at different height on the external surfaces of an apartment building. The positioning of the sensors has been accurately chosen to characterize the gradients of microclimatic parameters in the building area due to the different materials used, façade orientation, and the presence of natural features such as grass and trees. A crucial aspect in this pilot is that the data acquisition must occur at the same time for all nodes. For this reason, we decided to exploit the raw-LoRa protocol, which makes it easier to set a router as master and the end nodes as slaves, so that the master spreads its own time to the slaves making them synchronized before each measurement. The topology of the WSN is shown in Fig. 1, where a gateway (LoRa/wifi) serves as access point to the Web.

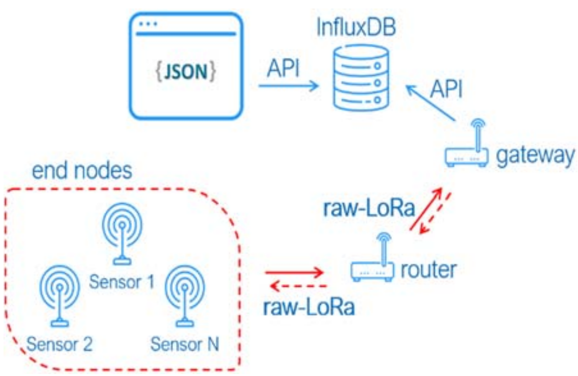


Fig. 1. raw-LoRa wireless sensor network set up for the SINFONIA project. The architecture encompasses 20 end nodes, a router, a gateway, and an InfluxDB database, whose API can be exploited to store and retrieve the data acquired by the sensors.

The LoRa packet management performed in the SINFONIA WSN is sketched in Fig. 2. The functioning has been obtained by programming each component of the network, namely the sensor nodes, the router and the gateway, in a stand-alone way. The operating routine starts with the router spreading its own real time clock (RTC) signal for 40 seconds. In the meantime, all the other system components, set to receiving mode, wait for a LoRa packet containing the router's RTC. Once the sensor nodes and the gateway are synced, the whole system goes into wait mode (machine idle) till a preset time. Then all the nodes acquire contemporary a measurement. That done, the nodes forward the data (temperature, relative humidity, battery level) to the router in a staggered way, i.e. one by one every 2 seconds, and then go into deep sleep mode.

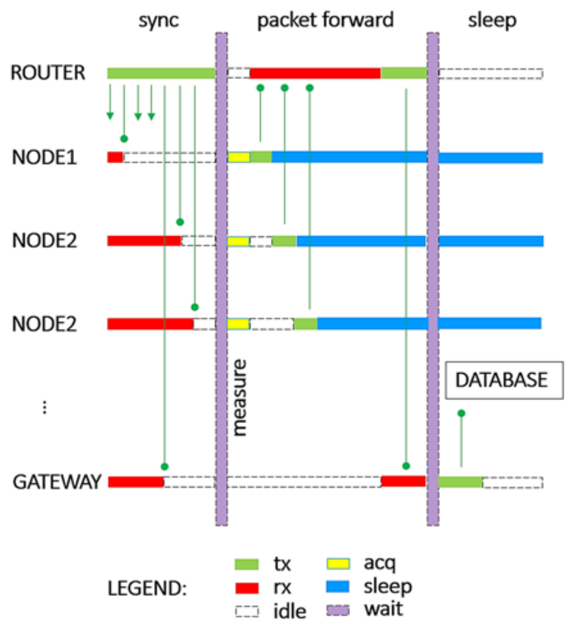


Fig. 2. Sketch of the working plan for the SINFONIA WSN, which repeats itself from left to right in an infinite loop every 10 minutes. The layers represent the main hardware components of the WSN, which are the router, the sensor nodes (only two are reported in the figure), and the gateway. The color code used aims at showing the different modes in which the components operate: green for transmitting mode, red for receiving mode, yellow for measuring, white for idle, cyan for sleeping.

After 60 seconds from the measurement time (violet vertical bar in the figure), the router wraps all the data received from the sensor nodes in a single LoRa packet and send it toward the gateway, which is still set to receiving mode. The gateway, in turn, posts the sensor data into the InfluxDB database [10] through a Wi-Fi connection via an HTTP POST request and goes into deep sleep, as well. The only component still on now is the router, which spreads the RTC signal every 1 minute in case of some nodes went out of synchronization during operation. After 9 minutes from the measurement time, the nodes and the gateway wake up, ready for a new synchronization, and the routine starts again.

3.2. LoRaWAN

In the context of the BEACON Südtirol - Alto Adige project another architecture has been put in place to grant LoRaWAN connectivity inside the technology park of Bolzano. The aim of the pilot was to prove whether an IoT testing network can be set up by using only open source components. LoRaWAN is particularly suited for this, as it can manage many different end nodes at once. Therefore, we exploited a plain LoRaServer installation for the network management, integrated with an InfluxDB instance (see Fig. 3).

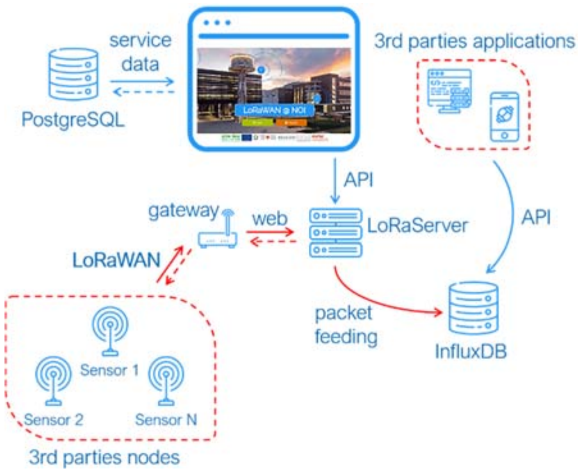


Fig. 3. Schematics of LoRaWAN network deployed at the NOI Techpark for BEACON project. The topology consists of several gateways that ensure indoor and outdoor connectivity in the Technology Park area. The network management is achieved through LoRaServer and a dedicated Web portal allows for the registration of new end nodes. Data transmitted over the network are stored into an InfluxDB.

The information gathered in the sensor-nodes is sent to the gateway over the LoRaWAN protocol. The gateway, in turn, is responsible for data forwarding via a User Datagram Protocol (UDP) link to the back-end, which is made of the LoRaServer and the InfluxDB components. As we wanted to keep the solution as much open as possible, we adopted a Raspberry Pi 3b+ as main board for the gateway and we achieved the LoRaWAN compliance through a RHF0M301 LoRa shield [11].

Several gateways have been deployed within the NOI Techpark area, which serve as access points to the LoRaWAN network. They are set up to have direct access to the network of the NOI Techpark (ScientificNetworkSouthTyrol [12]) through a secured virtual private network (VPN) connection at boot. This way, the gateways, which are assigned to a stack of static internet protocol (IP) addresses, can reach the

LoRaServer from any wireless local area network (WLAN). This strategy allows for a flexible deployment of the LoRaWAN access points, which can also be extended outside the NOI Techpark areal, at will.

In order to ensure the stability of the connection between the sensor-nodes and the backend, we took measures to make the gateways resilient and fault tolerant. To accomplish this goal, all the gateways have been provided with periodic checks to the status of the Wi-Fi and VPN connection and a fully automated routine makes them connect to the Wi-Fi network and to the VPN right after boot, in case of power outage. This design was created with the intention of keeping the downtime of the whole network low, which implies a low percentage of lost information. In such way, we achieve a network that, in case of faults, doesn’t require any human intervention to get restored.

We also set up a Web portal through which both custom and commercial sensor nodes can be added to the network without taking any subscription out. It acts as a unique access point to the LoRaServer and the InfluxDB application programming interface (API). By the one side, the LoRaServer API allows to register and manage the devices on network server. On the other side, the InfluxDB API can be used to retrieve the stored timeseries in a JSON format.

The portal (https://lorawan.beacon.bz.it/) is a Java Web Application running on Apache Tomcat [13], which implements registration, deletion, device status check, application and user management by means of servlets. These are defined as services inside the Web application to allow the Webpage to be a dynamic one.

Fig. 4 shows the administration page of the LoRaWAN Web portal. It has been developed in such a way to make network’s users autonomous in managing their own end nodes. After registration, the user is immediately redirected to the administration dashboard that encompasses five tabs: *Your devices, Add new application, Add new device, Get data, Settings*. At the first log-in, the devices list (first tab) is empty and the user needs to go through the creation of a new application (second tab) and the registration of new device (third tab).

Fig. 4. Admin Webpage of the LoRaWAN portal. The Web portal consists of a Java Web Application running on Apache Tomcat.

After doing so, a new line will appear in the devices list, which will carry the LoRaWAN device parameters in the format: *Device address, Application session key, Network session key*. If these three values match the ones of a working sensor node, the user will be able to retrieve the transmitted data through the InfluxDB API (fourth tab). Such an API can be consumed by Structured Query Language (SQL) through HTTP GET request. The result is a JSON carrying the requested timeseries. In the end, the settings page (fifth tab) allows to update/delete the user profile.

4. Sensor-nodes

To assess the performances of the two architectures, similar sensor-nodes have been added to both the raw-LoRa and the LoRaWAN networks (Fig. 5 - left).

The sensor node schematics is detailed in Fig. 5 - right. It encompasses a Pycom LoPy4 with expansion board version 3.0 as core unit. A DFrobot solar LiPo Charger, coupled with a 3.7 V – 5200 mAh battery and a 0.5W monocrystalline Seeed solar panel, manages the power supply. The sensing element is a Honeywell Humidicon HIH6121 enclosed into a plastic radiation shield and connected to the microcontroller via the I2C bus. The sensor node carries a microSD flash memory for local storage of acquired data. The system is completed by a Laird 868MHz 2 dBi antenna and a Bopla IP68 case.

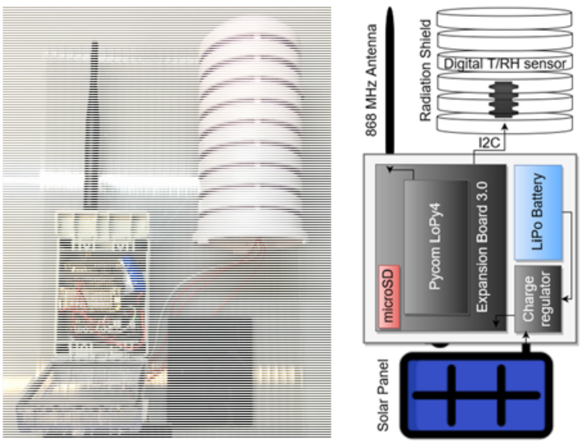


Fig. 5. Picture and schematics of one the sensor-node.

The sensor-nodes added to the WSNs share the same hardware, but they slightly differ in the programming script and data acquisition. In fact, for the raw-LoRa it is only needed to open a socket to start transmitting packets toward the gateway, while for the LoRaWAN a network session and an application session authentication must be performed to allow the data flow to the LoRaServer. Moreover, the sensor-nodes acquire different parameters. In the first case (raw-LoRa), the sensing element is inside a radiation shield so as to acquire outdoor temperature and relative humidity. In the second (LoRaWAN) the sensor is enclosed into the electronics box to investigate the hygro-thermal stress the micro-controller unit undergoes during operation.

However, these differences do not impact on the performance comparison we want to carry out, which will focus only on the communication reliability of the two WSNs being examined.

5. Experimental Results

5.1. Outdoor

Fig. 6 brings together some early timeseries from the SINFONIA's study-case. The charts have been obtained by means of Grafana [14]. The use of this analysis/visualization Web-tool is particularly suited to this purpose, as it comes with a native data source plugin for InfluxDB database and allow for automatic updates when new data are uploaded.

In the first three charts from top, temperature (red curves) and relative humidity (blue curves) from three sensor-nodes of the raw-LoRa WSN are reported. The data are acquired every 10 minutes and acquisition slot shown is 931 hours long. The daily oscillation of temperature and relative humidity are clearly visible in all plots, as well as the effect of a strong raining event in the middle of the timeseries. The difference between the three curves of temperature and the three curves of humidity is due to two main reasons. The first is sensor accuracy, which is 0.5 °C for temperature and 4 % for relative humidity. The second reason is because the sensor-nodes are located at different positions (mainly different heights) on the façades of district to be monitored.

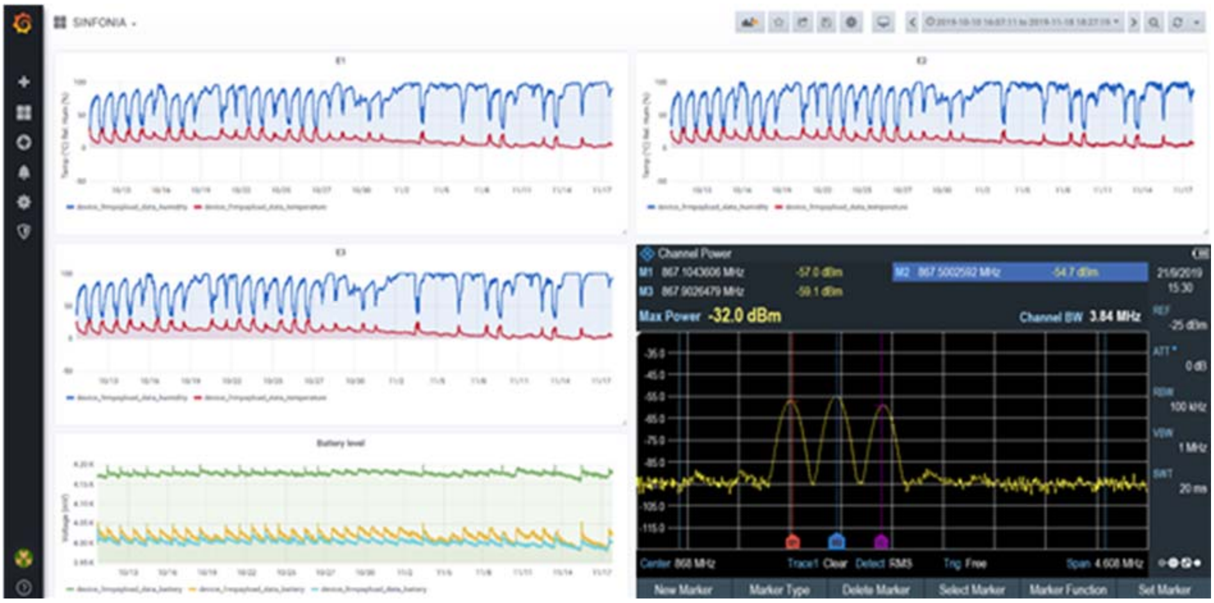


Fig. 6. Grafana dashboard created to monitor SINFONIA's campaign in real-time. The upper charts represent the temperature (red curve) and relative humidity (blue curve) acquired by four sensor-nodes placed on the façade of the building under analysis. The bottom-left chart shows the battery level of all sensor above, measured in residual voltage amplitude. In bottom-right, the screenshot of the frequency splitting on the LoRa channels is demonstrated when the three sensor-nodes transmit at the same time.

The bottom-left chart carries the battery level of all the sensor-nodes. The difference in the average level of the green curve in respect to the other ones is due to a better health status of that precise battery. The sawtooth behavior can be ascribed to the fact that the curves show the voltage amplitude of the batteries, which decreases steeply immediately after recharging phase has ended. Then the voltage undergoes a shallower decrease till the next charging peak, occurring during daytime. It would be easy to include some algorithms into the micro-controller script that convert the voltage in the residual capacity [15]. However, this falls out of scope for this work, as it would bring more complexity and more power consumption at node level, possibly lowering the reliability of the solution.

In the bottom-right of Fig. 6, the screenshot of an FPH Rodhe&Schwarz spectrum analyzer is added to show how the transmission of the different sensor-nodes occur over the LoRa channels. This condition where all the nodes transmit information at the same time is not common in the raw-LoRa architecture, as the packet forwarding is scheduled to be split in time among nodes. However, the measure demonstrates that unwanted emissions and leakage power form adjacent channels is negligible in such a system.

Permissible frequency deviation and RF waves emitted by the secondary will be detailed in a future work.

5.2. Indoor

Even if LoRa-based protocols are not particularly suited for indoor connectivity, mainly due to the long-wave nature, we tried to get the best out of this communication protocol within the spaces of the NOI Techpark.

Coverage and performance obtained with a single gateway onto a floor of more 500 m^2, with void space in the middle, are reported in Fig. 7. In order to estimate Received Signal Strength Intensity (RSSI) and Signal to Noise Ratio (SNR) in the different rooms, the figures of merit collected at the gateway have been analyzed. In fact, the LoRaWAN data model carries RSSI and SNR, as well, and the LoRaServer allows for an easy access those values. The maps shown in the figure have been obtained by acquiring multiple measurements per each room, where the gateway is kept steady into the upper-left room, and a 14dBm transmitting end-node is moved from space to space. The averaging over 10 measures per room demonstrate an RSSI ranging from -23 to -95 by passing from the gateway room to the most far apart location of the floor. Same happens for SNR, which varies from 12.8 to 4.2.

We want to remark that during the test, all the packets have been correctly delivered even if the building has steel-reinforced walls surrounding the void space in the middle.

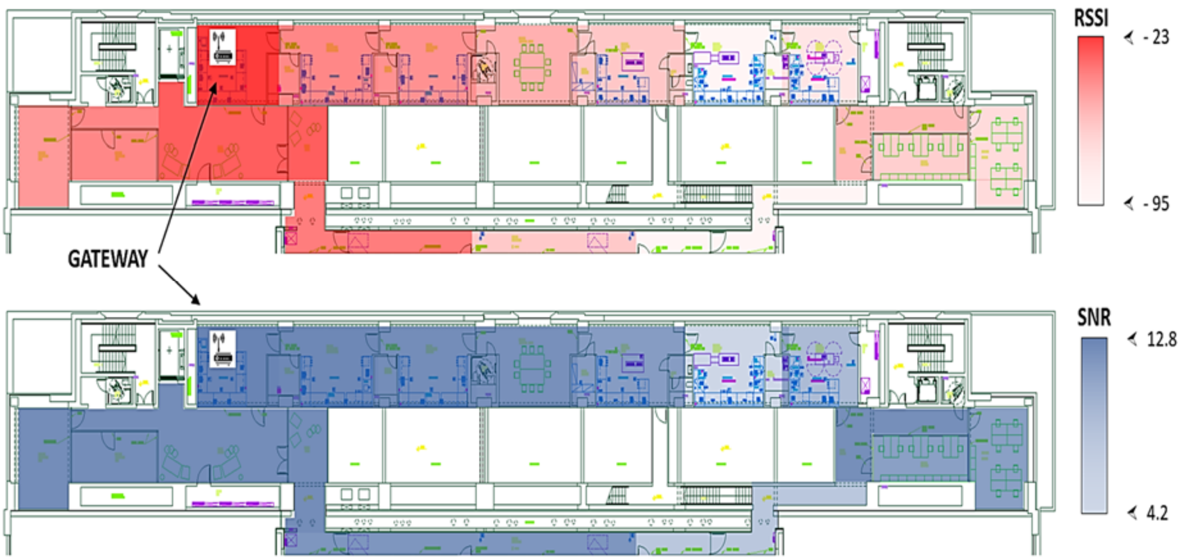


Fig. 7. Maps of Received Signal Strength Intensity (top) and Signal to Noise Ratio (bottom) superimposed to the layout of one of the NOI Techpark floors. The contour plots have been obtained by placing a 14dBm transmitting end-node in each room of the floor and averaging the stats data measured at the gateway.

5.3. Comparison of raw-LoRa and LoRaWAN

In Fig. 8, the two WSN architectures is compared in terms of reliability performances. The 160 hours long timeseries shown in the plots show the temperature in red solid line, the relative humidity in blue solid line (for a longer timeseries comparison see [16]). Despite the overall behavior is quite similar, the data from the two sensor nodes differ in range due to the different placement of the sensing element, which in the raw-LoRa node is within a radiation shield (outdoor sensor), while in the LoRaWAN node is inside the IP68 case (enclosed sensor). This aspect brings also to smoother trends of the temperature and the relative humidity in the second chart, because the sensor suffers from an enhanced sensing inertia [17] within the sealing case.

The figure insets aim to focus on some specific moments when lack of communication/connectivity has occurred during the acquisition. From the scatter plot we can see that network-failures happen to be randomly distributed. Moreover, a deeper data analysis allows us to claim that no communication lacks longer than 30 minutes (3 consecutive measurements) have occurred since the acquisition has started.

It is worth to note that this issue happens for both the WSNs in like manner, meaning that this is mainly related to interference/collisions at the LoRa frequencies [18-19]. Packet loss ratio figures for both the WSNs have been calculated over two 10000-sample timeseries. The obtained results are reported in Table 1.

The value reported in Table 1 for LoRaWAN is in agreement with previously reported analysis [20-21], as the node is less than 100 m far from the gateway.

Table 1. Packet loss ratios calculated on 10000 packets transmitted over the two WSN in comparison.

	raw-LoRa	LoRaWAN
Packet loss ratio	1.9 %	1.5 %

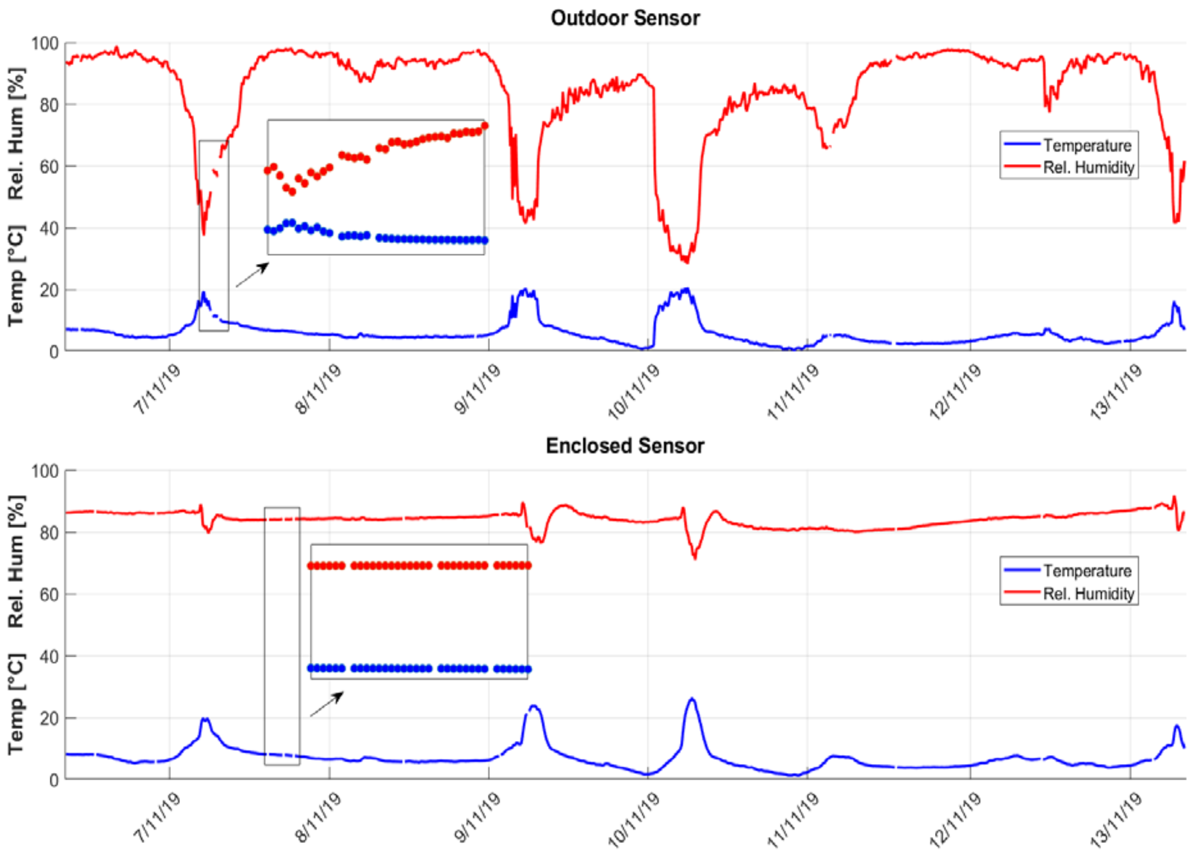


Fig. 8. 160 hours long timeseries acquired by one sensor node of the raw-LoRa network (top) and by one sensor node of the LoRaWAN network (bottom). Both the sensor nodes measure the temperature and relative humidity every 10 minutes. The lack of communication occurred in specific moments are highlighted in inset. The difference between the two charts comes from the fact that, despite the hardware is the same, the raw-LoRa node measures outdoor parameters while the LoRaWAN node is sensing temperature and relative humidity inside the IP68 case.

6. Conclusions

With this comparative analysis, we show that the raw-LoRa and the LoRaWAN protocols share similar reliability. However, the first one happens to be more appropriate for use cases where there is the need of specific requirements, like synchronous acquisitions or direct communication among devices. Instead, the second one is best suited to integrate many different data sources with enhanced information management and sharing capabilities.

A common feature to both the approaches described in this work is that they entirely rely on open-source software. This proves that reliable acquisition systems based on low-cost and vendor un-locked components are now within reach. In fact, the solutions presented in this work are easily replicable and can be modified to the need or upgraded with ease. Data sharing is also eased, as we put special attention to use unified and standard data model. In such way, we gain full control on data work-flow, from the sensor-nodes to the final dashboard for the end-user. Moreover, the proposed architectures are suitable for integration within wider digital platforms [22], and eventually they can be part of future developments toward a hybrid multi-protocol network for open IoT.

Acknowledgements

The research leading to these results has received funding by the European Union’s Seventh Framework Program FP7/2007–2013 Grant No 609019 and from the European Regional Development Fund (ERDF) Programme 2014-2020, under Project number FESR2023 (Beacon Sudtirol-Alto Adige) and Project number FESR1094 (Data Platform and Sensing Technologies for Environmental Sensing Lab).

References

[1]. N. Harris, J. S. Curry, Development and range testing of a LoRaWAN system in an urban environment, *International Journal of Electronics and Communication Engineering*, Vol. 12, Issue 1, 2018, pp. 43-51.

[2]. T. B. Ottosen, P. Kumar, Outlier detection and gap filling methodologies for low-cost air quality measurements, *Environmental Science: Processes & Impacts*, Issue 4, 2019.

[3]. A Jarwan, *et al.*, Data Transmission Reduction Schemes in WSNs for Efficient IoT Systems, *IEEE Journal on Selected Areas in Communications*, Vol. 37, Issue 6, 2019, pp. 1307-1324.

[4]. LoRa® Alliance Technical Marketing Workgroup, What is LoRaWAN, *White Paper*, 2105, p. 7.
[5]. SINFONIA (www.sinfonia-smartcities.eu).
[6]. G. Lobaccaro, S. Croce, D. Vettorato, S. Carlucci, A holistic approach to assess the exploitation of renewable energy sources for design interventions in the early design phases, *Energy Buildings*, Vol. 175, 2018, pp. 235-256.
[7]. S. Croce, D. Vettorato, R. Paparella, A Systemic Approach for the Optimization of Urban Surfaces Usage, in *Proceedings of the IOP Conference Series: Earth and Environmental Science*, Vol. 290, 2019, pp. 1-8.
[8]. BEACON BZ Webpage (https://beacon.bz.it/).
[9]. BEACON BZ - State of the art analysis, Whitepaper: The IoT Technologies (https://beacon.bz.it/wp-3/).
[10]. InfluxDB (https://www.influxdata.com/).
[11]. SeeedStudio LoRa/LoRaWAN Gateway Kit (http://wiki.seeedstudio.com/LoRa_LoRaWan_Gateway_Kit/).
[12]. ScientificNetworkSouthTyrol (https://www.sciencesouthtyrol.net/).
[13]. Apache Tomcat (http://tomcat.apache.org/).
[14]. Grafana (https://grafana.com/).
[15]. L. Chen, T. Laifa, F. Huanzhen. Li-ion battery capacity estimation: A geometrical approach, *Journal of Power Sources*, Vol. 261, 2014, pp. 141-147.
[16]. S. Tondini, raw-LoRa vs. LoRaWAN for Urban Monitoring: Two Wireless Sensor Network Architectures in Comparison, in *Proceedings of the Conference on Sensors and Electronic Instrumentation Advances (SEIA'2019)*, Tenerife (Canary Islands), Spain, 25-27 September 2019, pp. 265-269.
[17]. Y Xue, A. P. Cracknell, Advanced thermal inertia modelling, *International Journal of Remote Sensing*, Vol. 16, Issue 3, 1995, pp. 431-446.
[18]. G. Ferré, Collision and Packet Loss Analysis in a LoRaWAN Network, in *Proceedings of the 25th European Signal Processing Conference (EUSIPCO'17)*, Kos, Greece, 28 August-2 September 2017, pp. 2655-2659.
[19]. F. Van den Abeele, *et al.*, Scalability Analysis of Large-Scale LoRaWAN Networks in ns-3, *IEEE Internet of Things Journal*, Vol. 4, Issue 6, 2017, pp. 2186-2198.
[20]. J. Haxhibeqiri, *et al.*, LoRa indoor coverage and performance in an industrial environment: Case study, in *Proceedings of the 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'2017)*, Limassol, Cyprus, 12-15 September 2017.
[21]. B. Vejlgaard, *et al.*, Interference Impact on Coverage and Capacity for Low Power Wide Area IoT Networks, in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC'17)*, San Francisco, CA, USA, 19-22 March 2017.
[22]. A. Vianello, Environmental Data Platform (EDP), a Solution to Work with Big Data, Standardized and Reproducible, in *Proceedings of the Conference AGILE'19*, Limassol, Greece, 17-20 September 2019.

___

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 72-79

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Tunable Separation of Nanoparticles in a Continuous Flow Using Standing Surface Acoustic Wave

[1, *] Amar CHAALANE, [1] Mahmoud ADDOUCHE, [2] Rabah ZEGGARI, [1] Daniel GUNEYSU, [1] Franck LARDET-VIEUDRIN, [3] Amine BERMAK, [1] Céline ELIE-CAILLE, [1] Wilfrid BOIREAU and [1] Abdelkrim KHELIF

[1] FEMTO-ST Institute, UBFC, CNRS, ENSMM, UTBM,
15B Avenue des Montboucons, 25030 Besançon, France
[2] FEMTO Engineering, 15B Avenue des Montboucons, 25030 Besançon, France
[3] College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation,
Doha 5825, Qatar
[*] Tel.: +33(0)363082432, fax: +33(0)363082400
E-mail: amar.chaalane@femto-st.fr

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: Manipulating micro and nano-biological particles like extracellular vehicles (EVs), without extracting them from their biological media, presents a big challenge for diagnosis purposes. Here we present the design and fabrication of a sorting device based on the combination of microfluidic and electroacoustic modules that is capable of aligning and sorting submicron biological particles according to their size, compressibility or mass density, all in a tunable way. The device relies on a lithium niobate (LN) substrate to generate acoustic waves assembled with a micromachined glass layer for microfluidic circuits. The interference between the two surface acoustic waves (SAWs) generated by interdigitated transducers (IDTs) create a distribution of an acoustic radiation force (ARF). This force affects differently particles depending on their physical proprieties. The device is powered by an electronic circuit with a phase shifter to move the node of the standing surface acoustic wave (SSAW) along the channel width. When the device is powered at resonance frequency of the IDTs, experiment shows submicron particles alignment along the channel. By shifting the electrical signal between the two IDTs we can translate the pressure node at any targeted position in the channel width. The particles are then driven to one selected outlet.

Keywords: Nanotechnology, Microfluidic, Extracellular vesicles, Nanoparticles separation, Standing surface acoustic wave, Acoustic radiation force.

1. Introduction

The ability to sort, sieve, and manipulate small particles, especially in a tunable way, is highly required for various applications ranging from Micro-Nanoparticle sorting [1] to targeting drug delivery [2]. A conventional sieve uses a mesh of fixed size for separating smaller particles from the particle mixture of various sizes, but the limitation is that could not sort out different materials of the same size and is incapable of particle manipulation. Acoustic [3–8] and optical [9–13] tweezers have been developed to trap and manipulate nanoparticles [14-15], biomolecules [16-17], and organisms [18], yet they can capture only a few particles at a time and are difficult to scale up.

Otherwise, manipulating micro and nano-biological particles like extracellular vesicles (EVs), without extracting them from their biological media, presents a big challenge for the biological, clinical and pharmaceutical communities. Since EVs represents a highly important mediators in cell-to-cell communication [19–20], sorting and trapping them allows a better detection and qualification for diagnosis purposes.

Many technics have been used to improve the efficacy of particle sorting devices. In acoustic microfluidic devices, either bulk acoustic waves (BAWs) [21] or surface acoustic waves (SAWs) [22] are used to generate standing acoustic wave fields to trap particles realizing meso-scale particle patterning. This concept has become possible by the integration of acoustic and microfluidic systems in the same chip. It allows controlling the interaction between the acoustic wave and the submicron particles flowing in a channel of about tens microns wide.

In this work we introduce an acoustic sorting that can align, sort, and transfer a large number of particles in a liquid according to their physical proprieties (size, compressibility or mass density), all in a tunable manner. This device, based on the combination of microfluidic and electroacoustic systems, not only overcomes the shortfalls of a traditional sieve and acoustic or optical tweezers, but also fulfills many more functionalities.

The concept is motivated by the highly localized radiation force induced by the artificially engineered interdigitated (IDT) fingers patterned of a piezoelectric substrate to fulfill acoustic field. It is thus distinct from other acoustic techniques that rely on bulk standing waves or Gaussian beams directly generated by the acoustic transducer to trap one or several identical particles [3–8]. Such a resonance transmission and highly localized field in the region between the two IDTs induce a large acoustic radiation force that can be used to trap and sort small particles of few 100 nm of size. These techniques target to sort only one particle category (size range) at the same time. While in the device we will present in this paper, particles of different size can be gathered, at the same time, in several parallel lines where each line contains a distinct category in term of particle size.

2. Theoretical Analysis

When a particle in a viscous fluid interacts with a stationary elastic wave, two types of forces act on this particle: the acoustic force F_a and the Stokes forceF_{Stokes}. Since the acoustic force has a sinusoidal shape, it tends to move heavy and/or large size particles from the extremum towards the nodes. The Stokes force (so-called drag force) is a resistance force to the acoustic force.

The acoustic force F_a on a compressible and spherical particle can be expressed as [23]:

$$F_a = \frac{\pi^2 {P_a}^2 {R_p}^2}{3\lambda} \beta_f \, \phi(\beta,\rho) \sin\left(\frac{4\pi}{\lambda}x\right), \tag{1}$$

where P_a is the acoustic pressure, R_p is the particle radius, β_f is the compressibility of the fluid, $\phi(\beta,\rho)$ is the acoustic contrast factor, λ is the acoustic standing wave wavelength and x is the particle position across the fluidic channel ($x = 0$ in the center of channel). The acoustic pressure is further determined from the device characteristics:

$$P_a = \sqrt{\frac{PZ}{A}}, \tag{2}$$

where Z is the acoustic impedance of the substrate, A is the IDT area, and P is the power of the input signal. The mechanical properties of MVs are represented by the acoustic contrast factor:

$$\phi(\beta,\rho) = \left(\frac{5\rho_p - 2\rho_f}{2\rho_p + \rho_f}\right) - \left(\frac{\beta_p}{\beta_f}\right), \tag{3}$$

where ρ and β are the density and the compressibility of the particle, respectively. The indexes p and f corresponds to the particle and the fluid, respectively.

In the case of micro/nano-vesicles flowing in a fluid buffer, $\phi(\beta,\rho)$ will be positive. Therefore, vesicles move toward the acoustic force node positions.

According to the symmetry of the SAW system, the initial position of the pressure node (without applying a phase-shift) is normally located on the SAW system symmetry line. Therefore, this position corresponds to the center of the microfluidic channel. A phase-shift between the electrical signals of excitation of the IDTs makes it possible to shift the node to any desired position over the width of the channel, in a tunable manner.

The second force is the Stokes force F_{Stokes}:

$$F_{Stokes} = -6\pi\eta R_p\left(u_p - u_f\right), \tag{4}$$

where η is the dynamic viscosity of the fluid, R_p is the particle radius and u is the velocity. The indexes p and f corresponds to the particle and the fluid, respectively.

We obtain the motion characteristics of vesicles in a viscous flow by solving the equation:

$$F_a + F_{Stokes} = 0, \tag{5}$$

that gives:

$$\frac{\pi^2 PZ{R_p}^2\beta_f}{3\lambda A}\left[\left(\frac{5\rho_p - 2\rho_f}{2\rho_p + \rho_f}\right) - \left(\frac{\beta_p}{\beta_f}\right)\right]\sin\left(\frac{4\pi}{\lambda}x\right) - 6\pi\eta R_p\left(u_p - u_f\right) = 0 \tag{6}$$

The acoustic force is proportional to the particle volume and the drag force to the particle diameter.

This means that larger vesicles move faster to the pressure nodes, while smaller vesicles continue to move along the channel, at their original positions of entry into the acoustic opening, very little affected by acoustic force. Smaller particles will be aligned at the bellies (extremums) positions because the larger particles push them away of the nodes.

Indeed, the transit time τ of MVs moving from the channel center to the sheath flow is $\sim 1/R_p{}^2$, which enables size-selective vesicles separation [23]. This is why, in our design, we fixed the acoustic aperture at 5.2 mm to be larger than the distance traveled by the particles in a time duration τ. In this configuration, we can be sure that the particle motion reaches the steady state regime before coming out of the acoustic zone.

3. Methodology: a Tunable Manner for Nanoparticle Manipulation

The principal of particles manipulation using SAW is to generate a standing wave starting from two traveling surface acoustic waves (TSAW) that propagates in two opposite directions. This SSAW drives the particle toward one targeted position in the channel center that correspond to the pressure node (cf. Fig. 1).

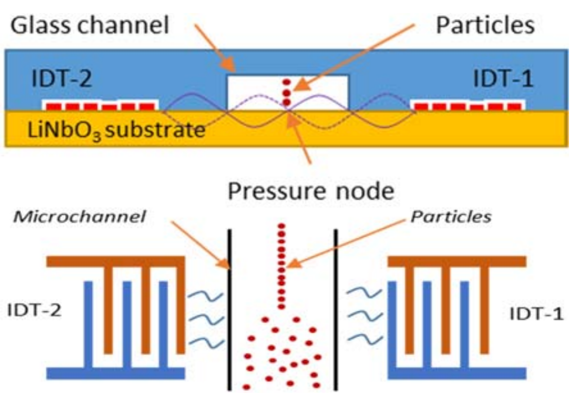
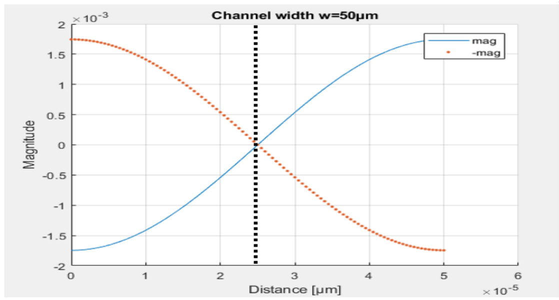


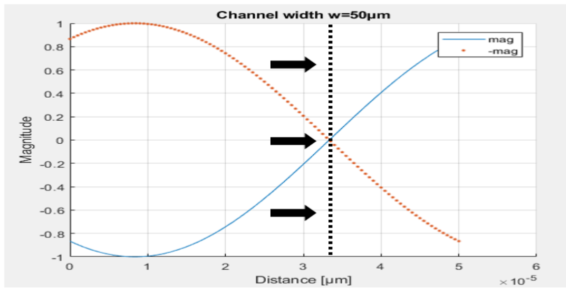
Fig. 1. Schematic of the acousto-fluidic device. It is composed of two IDTs patterned on piezoelectric substrate and a microchannel realized on glass substrate. The acoustic force drives larger vesicles to move faster to the pressure nodes, while smaller vesicles, very little affected, continue to move along the channel.

The standing wave occurred in the microchannel region generates an acoustic radiation force (ARF) that affects differently particles in fluid depending on their properties. The heavy and dense particles tend to tighten towards the nodes while the light and compressible particles remain on the antinodes [24-25]. This allow aligning the large size particles in the pressure node and the small size one in the anti-nodes.

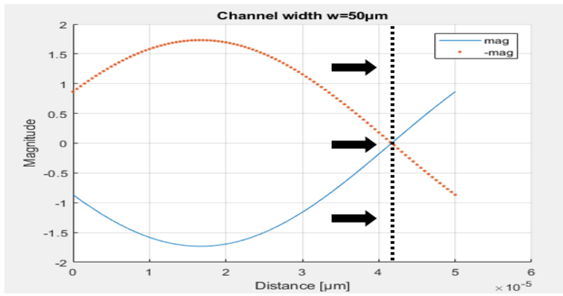
The idea here is to manipulate the position of the aligned particles simply by a continuous variation of the electrical signal phase between the two IDTs. This allows translating the node of the standing wave from the channel center to any targeted position situated in the channel width. As so, we can align and drive particles toward one desired channel outlet among the other outlets of the microfluidic circuit. Fig. 2 shows MATLAB software simulations of the electrical phase shifting effect on the position of pressure node, and this for fours relative phase shifting angles of 0°; 60°; 120° and 180°.



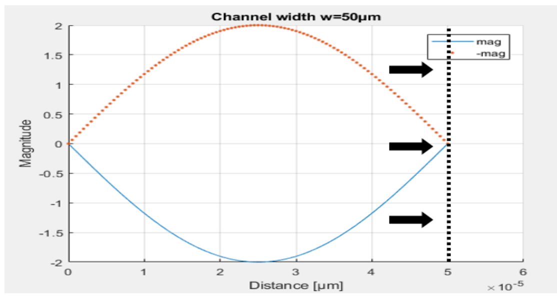
a) Δφ=0°



b) Δφ=60°



c) Δφ=120°



d) Δφ=180°

Fig. 2. Simulations of the tunable manner to translate the pressure node position from the center of the channel to the right wall, by shifting the electrical signal of 0°; 60°; 120° and 180°.

4. Experimental Results

4.1. Design and Geometrical Parameters

The design of the implemented acousto-fluidic system with the details of the microfluidic circuit and the network of fingers that compose the IDTs are presented in Fig. 3. It is composed of two layers of 3×3 cm^2 area: a glass layer for the microfluidic circuit and a piezoelectric layer for the acoustic wave generation.

The microfluidic circuit contains a main channel of l_c of length, w_c of width and h_c of depth, and six secondary channels. Each one of them ends on a circular hole that crosses the entire glass wafer thickness. Three holes are considered as "inlets" for the fluid injecting while the other three holes as "outlets" to recover either buffer or sorted particles.

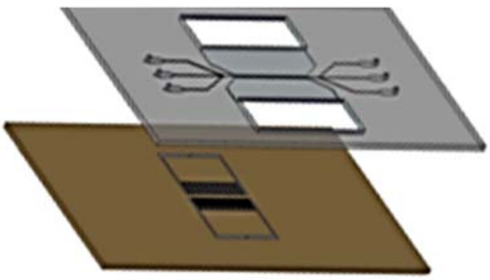


Fig. 3. A schematic view of the submicron particles separation micro-device with (top) the glass layer and (bottom) the piezoelectric layer. For the glass layer, we consider the three holes on the right side of as inlets while the other three holes as outlets for the fluidic.

The glass channel of 9 mm of length can be functionalized and coated to limit unwanted adsorptions. The use of glass allows a better biocompatibility, and a multi-use system. It is also to overcome limitation of PDMS for biology. Indeed, this elastomer ages, is sensitive to exposure at some chemicals and permeable enough to gas [26].

Two interdigital transducers (IDTs) are patterned on piezoelectric substrate of a lithium niobate ($LiNbO_3$). This material is very used for the manufacture of surface acoustic wave devices because of it high piezoelectric coupling factor. The two IDTs generate SSAW and thus acoustic force is created inside the channel. We fixed the width of the fluidic channel to 60 µm in order to match the acoustic half-wavelength. Based on the speed of sound of the SAW and the acoustic wavelength, the frequency of 40 MHz is used to supply the SAW transducers. In these conditions, the microchannel has one pressure node located at the center and two pressure anti-nodes located close to the walls. However, this position can be altered via a phase shift to any wanted position between the two IDTs.

The IDT electrode geometry is shown in Fig. 4, where the single-electrode-type (so-called solid-electrode-type) configuration has been adopted for the excitation of SAWs. The single-electrode-type IDT is widely used because of its structural simplicity and relatively wide strip width ($\sim \lambda/4$), which reduces the required resolution of photolithography [27].

The width w_e and pitch p_e of IDT electrodes are all 25 µm. This configuration generates SAW with wavelength λ of 100 µm. The distance between IDTs L is 700 µm. Because the behavior of SAW excitation at the IDT limit is somewhat different from that in its center region, the acoustic aperture W have to be large (here W = 5.2 mm); this is to avoid a significant diffraction effect of the traveling wave that take place in the IDT edge [27]. Furthermore, a larger aperture let a long nanoparticles-SSAW interaction time, allowing high flow particle sorting. The number of IDT finger-pairs (N=20) must be large enough to generate acoustic power capable of interacting with the nanoparticles but not very large to make the characteristics of IDT complex. This complexity is caused by the reflection of Bragg which occurs when $\lambda \sim p_e$ corresponding to the resonance condition for SAW excitation [27].

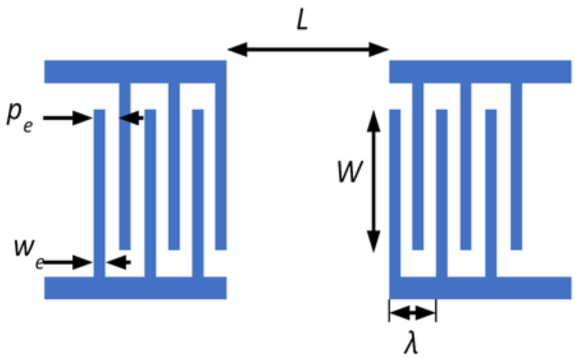


Fig. 4. Scheme of the SAW interdigitated transducer with the parameters of the electrode fingers.

4.2. Summary of the Fabrication Process

The acousto-fluidic device microfabrication process is mainly composed of three steps:

1) The realization of IDTs electrode onto a 128° Y-cut $LiNb0_3$ piezoelectric substrate.

2) The micro-machining of microfluidic circuit on a glass wafer: the wafer top side was micro-machined by deep reactive ion etching (DRIE) to achieve microchannels design. A femtosecond (FS) laser boring technic is used on the wafer backside to achieve holes of 500 µm in diameter for the fluidic inlets and outlets.

3) Both lithium niobate and glass wafers were stacked together in order to close the microfluidic circuit and to allow fluidic and acoustic interaction.

The realized SAW interdigitated transducer is presented in Fig. 5. It is composed of IDT-1 on the top and IDT-2 on the bottom, separates by the fluidic microchannel. The first wave propagates from top to down while the second one propagates in the opposite direction to fulfill a standing acoustic wave.

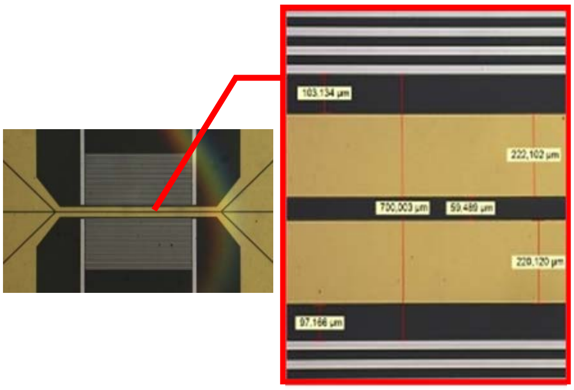


Fig. 5. The picture of the SAW interdigitated transducer that actuates with nanoparticles presents in a liquid. The last one is flowing in microchannel deposited on the $LiNbO_3$ substrate. The first wave propagates from top to down. The second one propagates in the opposite direction. The acoustic aperture is of 5.2 mm and the IDTs length is of 700 µm.

The microfluidic circuit is dry etched through a single step of photolithography until 30 µm of depth. Thus, the effective section of the channel is of trapezoidal shape with w_1=59.75 µm of width on the surface level and w_2=32.17 µm of width at the bottom level of the channel. Fig. 6 shows picture of the micro-acousto-fluidic device for nanoparticles manipulating and sorting.

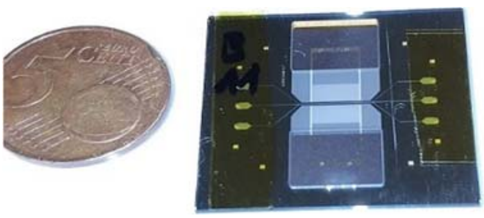


Fig. 6. A picture of the acousto-fluidic nanoparticles manipulator device. It is composed of a glass layer for the microfluidic circuit stacked to a lithium niobate layer for the acoustic standing wave.

4.3. Electrical Measurements

In order to characterize the microsystem, we have manufactured a mechanical support (Fig. 7) allowing to (1) host the device, (2) ensure the six fluidic connections to a controlled injection system and (3) connect IDTS for both the electrical characterization and SAW generation.

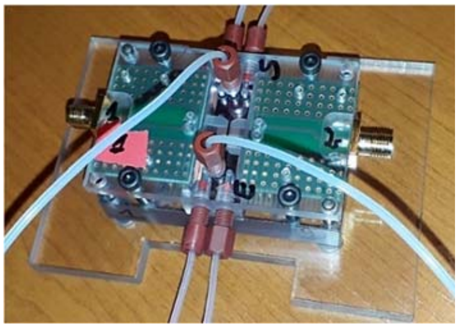


Fig. 7. The realized acousto-fluidic device with electric and microfluidic connections.

In Fig. 8 we represent a schematic view of the experimental setup. The device is powered by an AC sinusoidal signal generated by a, Anritsu 68147C synthesized signal generator. A power splitter divides the signal in two ways (0° and 180°) to supply the two IDTs. The first line (upper one) is used for the excitation of IDT-1. It contains an analog phase shifter to create a phase shift between the 2 IDTs electrical signal input. This allows to continuously tuning the SSAW pressure node along the channel width. Whereas the second line (downer one) is used for the excitation of IDT-2. Since the phase-shift response is not linear, the electric signal amplitude in the IDT-1 input might be different from those for IDT-2. That is why we use an attenuator in the bottom line to balance the electrical signal amplitude at the inputs of both IDT-1 and IDT-2.

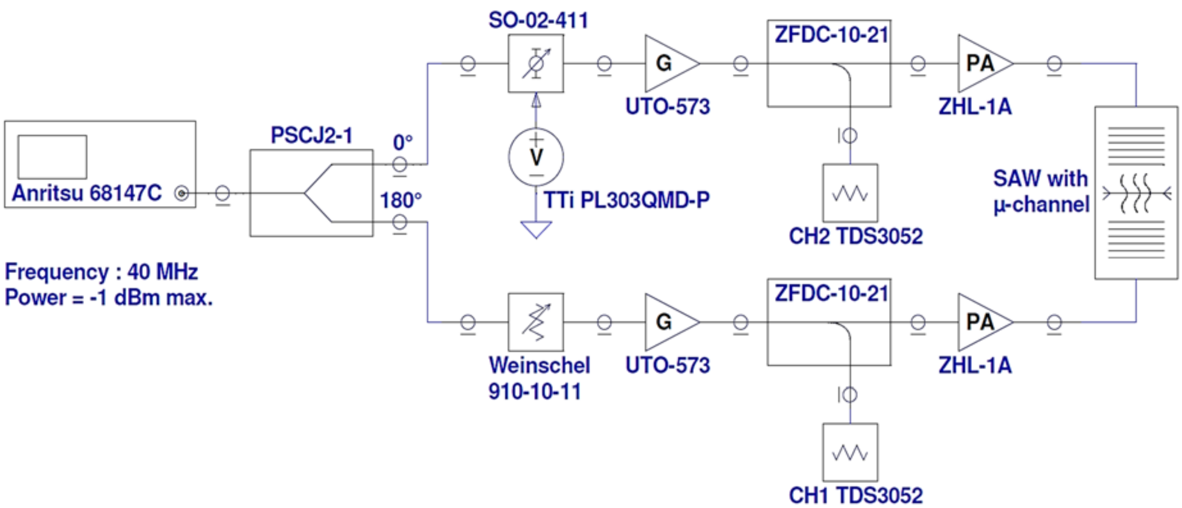


Fig. 8. Electronic circuit for the excitation of SAW interdigital transducer for the submicron acousto-fluidic sorting device.

For radio-frequency (RF) circuit analysis, the scattering parameters is widely used. The network analyzer commonly used for SAW device characterization. In fact, this makes it possible to measure insertion losses, rejection, bandwidth and the frequency of resonance of the SAW system.

Fig. 9 shows the insertion loss and reflection coefficient as a function of the frequency under these conditions. We can see that the central frequency: f_0 is 39.8 MHz which is very close to the simulation (40 MHz). The bandwidth is about 2 MHz, the loss of insertion is 30 dB and the rejection is 15 dB. Even if those parameters are not optimal, we were able to generate SSAW capable to drive submicron particles.

When the SAW system resonance frequency is determined, we can proceed with the optimization of the circuit by adding an electronic stage for an impedance matching of the IDTs by using the smith chart. The impedance between the RF source to the IDT electrodes need to be matched to maximize the electrical-mechanical energy transfer.

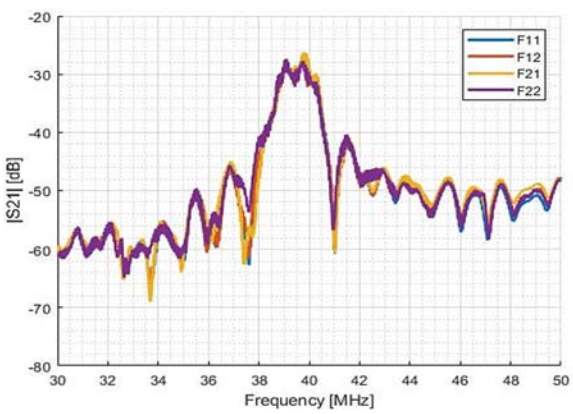


Fig. 9. Frequency response of the SAW interdigital transducer onto a 128° Y-cut $LiNb0_3$ substrate (for four identical devices F11, F12, F21 and F22). The central frequency f_0 = 39.8 MHz. The bandwidth is about 2 MHz, the loss of insertion is 30 dB and the rejection is 15 dB.

4.4. Microfluidic Measurements

In the following experiment, fluorescent calibration beads (480 nm in diameter/polymer) will be used to show the influence of surface stationary wave on submicron particles. We used a submicron synthetized beads from Micro particles GMBH®: MF-FITC-COOH, fluorescent beads (abs/em: 506/529 nm) with the diameter of 480 nm. These beads present carboxylic acid function allowing protein grafting on their surface, which permit to perform immunocapture, after sorting and recovery of particles.

Before introducing beads in microchannel, several conditioning steps are performed in the device: we wet the channel by using a water/methanol solution (50/50). The running buffer was PBS solution (Phosphate buffered saline) with Tween™ 20 (0.1 % v/v) from Thermo Fischer for a better wettability. To control the flow, the MFCS-EZ system from Fluigent® was used. The system allows the injection of sample in the microfluidic circuit depending on the pressure, from 0 to 1 bar but the working pressure was around 20 mbar. The experiment was monitored in real time by using a fluorescent microscope (AXIO SIP61023, ZEISS®).

5. Discussion

Fig. 10 presents a first analysis showing a laminar flow in the microchannel.

Initially, when we introduce nanoparticles in the inlet 1, it is automatically that the all particles exit from outlet 1. The beads are randomly dispersed along the half-channel in dynamic fluidic mode [28].

On the other hand, when we activate SAW with a normally phase–shifting of 180° between the IDTs, the randomly dispersed beads are aligned in a narrow line along the channel. This line is situated in the middle of the channel width which correspond to the pressure node position in the normally case (without supplementary phase-shifting) (cf. Fig. 11).

By shifting the SSAW phase with a relative phase shifting ranging from 0 to 180°, we translate beads from outlet 2 to outlet 1 or 3 according to the phase-shift we apply. As so, we realize sorting by driving 480 nm in diameter particles to one of the three channel outlets (Fig. 11). This is the first result showing that we can control vesicles flow and destination in specific outlet, just by acting on the phase of SSAW, to drive them from one side to other side.

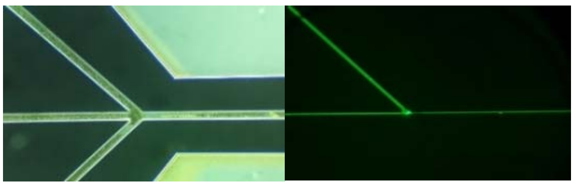


Fig. 10. Laminar flow experimentation using 480 nm in diameter beads in a buffer flowing throw a 60 µm of width fluidic channel: (left) optical view and (right) fluorescence microscopy images.

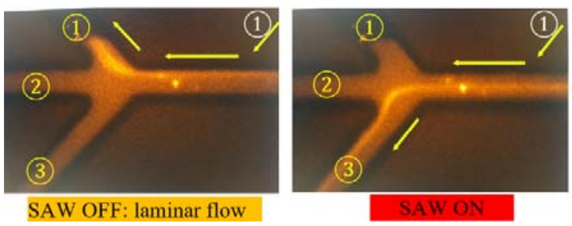


Fig. 11. Nanoparticles manipulation results using nanoparticles of 480 nm in the inlet 1. When we apply the SAW, we can move the particles from outlet 1 to outlet 2 or to outlet 3 according to the phase-shift we apply.

6. Conclusion

The main idea of this project is to be able to manipulate and sort the nanoscale vesicles by acoustic waves.

In this work, we experimentally validate the submicron particles alignment and sorting using SSAW microsystem. Experience has shown a submicron particles alignment at a fixed position along the channel. By shifting the SSAW phase, we can drive particles to one of the three outlets.

We have successfully fabricate an acoustic-based glass microfluidic system for label-free and continuous sorting of NP. We experimentally validated the particle manipulation and sorting using one type of 480 nm particles in glass microchannels. We validate the new concept of standing wave monitoring using phase-shifting.

By controlling the radiation force, a large number of particles can be manipulated in the same time. In fact, if a multiple particle sizes are presented in a continuous flow, each particle species of same size could be driven to one of the multiple microfluidic outlets.

Acknowledgements

This work is funded in part by NPRP grant # NPRP10-0201-170315 from the Qatar National Research Fund (a member of Qatar Foundation) and through the EIPHI Graduate School (contract "ANR-17-EURE-0002"). The authors also acknowledge the support of FEMTO-ST technological and characterization facilities. The findings herein reflect the work, and are solely the responsibility of the authors.

References

[1]. WU Mengxi, OZCELIK Adem, RUFO Joseph, *et al.*, Acoustofluidic separation of cells and particles, *Nature Microsystems & Nanoengineering*, Vol. 5, Issue 1, 2019, p. 32.

[2]. Ferrara Katherine, Pollard Rachel, Borden Mark, Ultrasound microbubble contrast agents: fundamentals and application to gene and drug delivery, *Annu. Rev. Biomed. Eng.*, Vol. 9, 2007, pp. 415-447.

[3]. WU Junru, Acoustical tweezers, *The Journal of the Acoustical Society of America*, Vol. 89, Issue 5, 1991, pp. 2140-2143.

[4]. BLACK Justin P., WHITE Richard M., GRATE Jay W., Microsphere capture and perfusion in microchannels using flexural plate wave structures, in *Proceedings of the IEEE Ultrasonics Symposium, Munich,* Vol. 1, 2002, pp. 475-479.

[5]. HAAKE Albrecht, NEILD Adrian, RADZIWILL Gerald, *et al.*, Positioning, displacement, and localization of cells using ultrasonic forces. *Biotechnology and Bioengineering,* Vol. 92, Issue 1, 2005, pp. 8-14.

[6]. LEE Jungwoo, TEH Shia-Yen, LEE Abraham, *et al.*, Single beam acoustic trapping, *Applied Physics Letters*, Vol. 95, Issue 7, 2009, p. 073701.

[7]. SHI Jinjie, AHMED Daniel, MAO Xiaole, *et al.*, Acoustic tweezers: patterning cells and microparticles using standing surface acoustic waves (SSAW), *Lab on a Chip,* Vol. 9, Issue 20, 2009, pp. 2890-2895.

[8]. LENSHOF Andreas, MAGNUSSON Cecilia, LAURELL Thomas, Acoustofluidics 8: Applications of acoustophoresis in continuous flow microsystems, *Lab on a Chip*, Vol. 12, Issue 7, 2012, pp. 1210-1223.

[9]. ASHKIN Arthur, DZIEDZIC James M., Optical trapping and manipulation of viruses and bacteria, *Science,* Vol. 235, Issue 4795, 1987, pp. 1517-1520.

[10]. MACDONALD Michael P., PATERSON Lynn, VOLKE-SEPULVEDA K., *et al.*, Creation and manipulation of three-dimensional optically trapped structures, *Science,* Vol. 296, Issue 5570, 2002, pp. 1101-1103.

[11]. FAZAL Furqan M., BLOCK Steven M., Optical tweezers study life under tension, *Nature Photonics*, Vol. 5, Issue 6, 2011, pp. 318-321.

[12]. THOMPSON Jeffrey Douglas, TIECKE T. G., DE LEON Nathalie P., *et al.*, Coupling a single trapped atom to a nanoscale optical cavity, *Science,* 2013, Vol. 340, Issue 6137, pp. 1202-1205.

[13]. ZHONG Min-Cheng, WEI Xun-Bin, ZHOU Jin-Hua, *et al.*, Trapping red blood cells in living animals using optical tweezers, *Nature Communications*, Vol. 4, 2013, p. 1768.

[14]. JUAN Mathieu L., GORDON Reuven, PANG Yuanjie, *et al.*, Self-induced back-action optical trapping of dielectric nanoparticles, *Nature Physics*, Vol. 5, Issue 12, 2009, p. 915-919.

[15]. MANDAL Sudeep, SEREY Xavier, ERICKSON David, Nanomanipulation using silicon photonic crystal resonators, *Nano Letters*, Vol. 10, Issue 1, 2009, pp. 99-104.

[16]. PANG Yuanjie, GORDON Reuven, Optical trapping of a single protein, *Nano Letters*, Vol. 12, Issue 1, 2011, pp. 402-406.

[17]. YANG Allen HJ, MOORE Sean D., SCHMIDT Bradley S., *et al.*, Optical manipulation of nanoparticles and biomolecules in sub-wavelength slot waveguides, *Nature*, Vol. 457, Issue 7225, 2009, pp. 71-75.

[18]. DING Xiaoyun, LIN Sz-Chin Steven, KIRALY Brian, *et al.*, On-chip manipulation of single microparticles, cells, and organisms using surface acoustic waves, in *Proceedings of the National Academy of Sciences,* Vol. 109, Issue 28, 2012, p. 11105-11109.

[19]. SONG Dongli, YANG Dawei, POWELL Charles A., *et al.*, Cell–cell communication: old mystery and new opportunity, *Cell Biology and Toxicology,* Vol. 35, Issue 2, 2019, pp. 89-93.

[20]. HUANG-DORAN Isabel, ZHANG Chen-Yu, VIDAL-PUIG Antonio, Extracellular vesicles: novel mediators of cell communication in metabolic disease, *Trends in Endocrinology & Metabolism*, Vol. 28, Issue 1, 2017, pp. 3-18.

[21]. HAAKE Albrecht, NEILD Adrian, RADZIWILL Gerald, *et al.*, Positioning, displacement, and localization of cells using ultrasonic forces, *Biotechnology and Bioengineering*, Vol. 92, Issue 1, 2005, pp. 8-14.

[22]. COLLINS David J., MORAHAN Belinda, GARCIA-BUSTOS Jose, *et al.*, Two-dimensional single-cell patterning with one cell per well driven by surface

acoustic waves, *Nature Communications*, Vol. 6, 2015, p. 8686.

[23]. LEE Kyungheon, SHAO Huilin, WEISSLEDER Ralph, *et al*., Acoustic purification of extracellular microvesicles, *ACS Nano*, Vol. 9, Issue 3, 2015, pp. 2321-2327.

[24]. GRÖSCHL Martin, Ultrasonic separation of suspended particles - Part I: Fundamentals, *Acta Acustica United with Acustica*, Vol. 84, Issue 3, 1998, pp. 432-447.

[25]. JOHANSSON Linda, ENLUND Johannes, JOHANSSON Stefan, *et al*., Surface acoustic wave induced particle manipulation in a PDMS channel—principle concepts for continuous flow applications, *Biomedical Microdevices*, Vol. 14, Issue 2, 2012, pp. 279-289.

[26]. CHARATI S. G., STERN S. A., Diffusion of gases in silicone polymers: molecular dynamics simulations, *Macromolecules*, Vol. 31, Issue 16, 1998, pp. 5529-5535.

[27]. HASHIMOTO Ken-Ya, HASHIMOTO Ken-Ya, Surface acoustic wave devices in telecommunications, *Springer-Verlag Berlin Heidelberg*, 2000.

[28]. A. Chaalane, D. Guneysu, M. Addouche, R. Zeggari, F. Lardet-Vieudrin, C. Elie-Caille, W. Boireau, A. Khelif, Continuous sorting of submicron particles in a pre-analytical device based on acousto-fluidic microsystem, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'2019)*, Tenerife, (Canary Islands), Spain, September 2019, pp. 25-27.

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 80-86

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

An Approximation to the Relationship Between Climatic Variables Obtained Through Remote Satellite Sensors and Hospital Admissions: A Case Study on Gran Canaria Island

[1] I. CABALLERO-LEIVA, [1] N. MARRERO BETANCORT, [2] J. J. RODRIGUEZ BETANCOR, [1, *] D. RODRIGUEZ-ESPARRAGÓN and [1] J. MARCELLO

[1] Instituto de Oceanografía y Cambio Global, IOCAG. Universidad de Las Palmas de Gran Canaria ULPGC, Campus de Tafira, 35017 Las Palmas, Spain
[2] Gran Canaria Dr. Negrín University Hospital, 35010 Las Palmas, Spain
[1] Tel.: +34 928457357, fax: +34 928452716
E-mail: dionisio.rodriguez@ulpgc.es

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: The risk to human health is among the possible consequences of climate change or global warming. In this sense, it is necessary to deepen between climate and health interactions, in order to establish scenarios and policies to mitigate their consequences. The Sahara dust is well known to have an adverse impact on human health. The Canary Islands, due to their vicinity to the Sahara Desert, are frequently affected by the Sahara aerosols carried by the winds. This study analyses the correlation between climatic (Sea Surface Temperature, Aerosol Optical Thickness, Wind Intensity and Wind Direction) and clinical (emergency admissions due to respiratory pathologies) variables in the island of Gran Canaria and for a period of 14 years. The bivariate lineal correlation (Pearson) shows statistical significance between accumulated monthly averaged values of Sea Surface Temperature and the emergency admissions. There is also a significant correlation between the climatic variables Wind Intensity and Aerosol Optical Thickness.

Keywords: Remote sensing, Health risks, Climate Change, Global warming, SST, AOT, Wind intensity, Wind direction.

1. Introduction

Climate change or global warming is one of the scientific challenges with the greatest social impact today. The number of papers published as well as the impacts in all types of media confirm this statement [1-2].

The main consequence of warming the Earth's surface is a change in the climate associated with adverse meteorological events that will occur progressively more frequently; extreme heat waves, natural disasters and rainfall variation [3-4]. Consequently, climate change influences the social and environmental determinants of health, namely clean air, clean water, enough food and safe housing [5].

Thus, to help establish policies to mitigate the effects associated with global warming, it is necessary to study the relationship between health and climate [6-8]; especially in those particularly vulnerable territories: islands and archipelagoes, and less developed regions [9].

On the other hand, the availability of data captured with satellite edge remote sensing sensors has increased significantly in recent years [10]. Likewise, the spatial, spectral and temporal resolution of the products they provide [11]. As a result, these technologies are used for the monitoring of the terrestrial surface in order, among others, to analyze climate change, as well as for the generation of time series of climate variables that allow to analyze their evolution and help in predicting future scenarios.

In this context, the main contribution of this work is to try to analyze the relationship between the evolution of climate variables obtained by remote satellite remote sensing and health, through hospital admission data in the environment of Canary archipelago. This work is a review of the one presented at Fifth International Conference on Sensors and Electronic Instrumentation Advances (SEIA' 2019) [12].

Canary Islands are frequently influenced by the Sahara dust that crosses the Atlantic Ocean sweeping the Spanish archipelago. This meteorological event is commonly known as "CALIMA" (Cloud, Aerosols and Ice Measurements in the Sahara Air Layer). This paper is part of a wider research project which studies the relationship between the CALIMA and the weather conditions and their influence on respiratory, psychic and allergic conditions.

The influence of Sahara dust in human health is well studied not only in several European countries [13-18] but also across the Atlantic Ocean, all the way in USA [19]. The emergency services of the Canary Islands are very familiar with the effect that this event has in the population´s health, noticing an increase in the attendance of patients during the days when the event is present (Doctor Negrín University Hospital, admissions team personal comment).

Sahara dust is transported along several paths around the globe; over the North Atlantic Ocean to North America and South America, across the Mediterranean to Southern Europe and sometimes as far north as Scandinavia or east as Middle East [20]. Considering the proximity of the Macaronesian Region to the desert´s coast and that the dust sweeps along the Canary Island when the wind blows towards the Atlantic, the relevance of studding the impact of these aerosols in the canarian population seems evident. Furthermore, there are studies looking into the possibility of the Trend Winds rotation to the east [21] which would increase the periods of time in which the canarian archipelago is exposed to the Sahara dust. Therefore, it is crucial for the medical services of the Canary Islands to understand what the consequences of such change in the wind climate could mean in the health of their residents in order to plan the adaptation of their services.

The project counts with the Admission Department of one of the main hospital in the island of Gran Canaria (Doctor Negrín University Hospital). They provide not only the clinical data but also their medical expertise to the study.

The section of the study presented in this paper shows results of the early stages of the project and focuses on the analysis of climatic and clinical parameters, their correlation and the influence on respiratory conditions in the gran canarian population.

The rest of the paper is organized as follows: in the next section the data as well as the methodology applied to analyze them are described; in Section 3 the results are presented and discussed; finally, in Section 4, the conclusions are detailed.

2. Data and Methodology

2.1. Study Area

The Canary Islands are in the west coast of the African Continent and are part of the Macaronesian Region. Its climate is characterized by the predominance of the Trade Winds, which reach the islands with a north-northeasterly direction [14], [22]. The wind climate of the islands is driven by the Azores high pressure system which dominate below approx. 1500 m a.s.l. [23]. When the dominant wind direction changes to easterly and southerly the islands receive Sahara air, which is characterized by warm temperatures, high relative humidity and the presence of suspended Sahara dust (CALIMA) [13]. This happens most frequently during the months of summer, July to August [15].

Although this phenomenon affects all the islands that constitute the archipelago, in this work we focus on the island of Gran Canaria. Gran Canaria offers a unique orography with an area of 1560.1 km^2 and a maximum altitude of 1956 meters [24]. Its population is 846,717 inhabitants, residing in the capital, Las Palmas de Gran Canaria, 383,343 people, making it the most populous city in the Canary Islands [25]. Specifically, as it will be seen in the section describing the clinical data, the study area will comprise the northern area of the island and the city of Las Palmas.

2.2. Data

The following subsections describe the medical and climate data that have been used in this work. According to its availability, the data cover the period from January 2003 to December 2017, 14 years in total.

2.2.1. Climatic Data

The following climate variables were selected for this study: Wind Direction (WD), Wind Intensity (WI), Sea Surface Temperature (SST) and Optical Aerosol Thickness (AOT), all expressed in monthly averages. The selection criterion has been its known influence on human health [26]. Indeed, the sea acts as a modeler of global temperature in an island environment of small dimensions, such as the island of Gran Canaria. Similarly, AOT, WD and WI act as elements that may condition the quality of the breathed air.

Wind data was obtained from the National Centre of Atmospheric Prediction and the National Centre for Atmospheric Research of the United States of America (NCEP/NCAR). They consist on a re-analysis of monthly averages of the u and v components of the wind measured at 10 meters above sea surface, both near shore and offshore. The data is provided in a matrix of 192 pixels of longitude and 94 pixels of latitude (Gaussian grid), with a processing level 4, and a geographical resolution of 0.25º × 0.25º (approx. 25 km × 25 km). It is provided in HDF5 format [27]. From the Gaussian grid, 8 time series were extracted around the archipelago. After comparison of WD and WI for all points and corroboration that there was no significant spatial variation of the wind data in the study area, an average was performed and so one single (averaged) time series was used for the analysis performed in the present study. Fig. 1 shows the location of the 8 NCEP extracted data points.

Fig. 1. NCEP data points locations. P1 to P8. (Google earth Pro. Image Landsat/Copernicus).

AOT [28] and SST [29] are a product of MODIS (Moderate Resolution Imaging Spectroradiometer), aboard the NASA satellite Aqua and presented in monthly averages. The data is provided in maps, the 4km resolution was selected for the tasks performed in this study. Six data points (time series) were extracted from the mentioned map. An average of the 6 data points was performed and so single averaged time series of AOT and SST are presented in this work. The AOT used is AOT_869. Fig. 2 shows the location of the 6 AOT and SST extracted data points.

Fig. 3 to Fig. 5 display the monthly time series of climate data for the entire study-period. Fig. 1 shows some cyclic pattern in the wind direction and wind intensity components. This responds to the predominant presence of trade winds blowing over the archipelago. Similarly, Fig. 3 shows the cyclicity of SST between cold and warm seasons. On the contrary, Fig. 2 shows the random pattern of the presence of aerosols on the atmosphere.

Fig. 2. MODIS-Aqua data points location. P1-P6. (Google Earth Pro. Image Landsat/Copernicus).

2.2.2. Clinical Data

The clinical data was provided by the admissions team of the Gran Canarian Doctor Negrín University Hospital. This center is one of the largest hospital facilities in the Canary Islands. Specifically, its area of influence encompasses the entire north of the island and the northern half of the city of Las Palmas de Gran Canaria. These facts were key criteria to the selection as reference center for this study.

The data consists on the number of daily emergency admissions due to breathing pathologies (IUPR). Since the climatic data was obtained in monthly averages, the daily clinical data was accumulated into the same time step.

The data was converted to 1000 %, for a population of 350000 citizens that were assigned to the Hospital when the study took place [30].

Fig. 6 displays the monthly clinical data in rates per thousand and for the entire study-period.

3. Results

The data analysis presented in this article consists on bivariate linear correlations and were applied to all pairs of variables. The Pearson coefficient was calculated for all pairs of climatic variables as well as between the climatic and clinical variables. Other than correlation analysis applied to the whole data series as one, seasonal analysis was also performed, as well as annual trends.

The following figures show some of the results obtained.

Fig. 7 shows the accumulated monthly averaged values of SST vs IUPR for the 14-year period, that is, the annual trends. The trend lines show an inverse relationship between these two variables. For the first trimester of the year, both variables see a progressive decrease, but while IUPR keeps dropping until the end of the year (with excuse of small raises in June and

September), the Sea Surface Temperature increases from March to September, to drop again from then to December. The temperature ranges between a maximum of 25º seen in September and 20º in March.

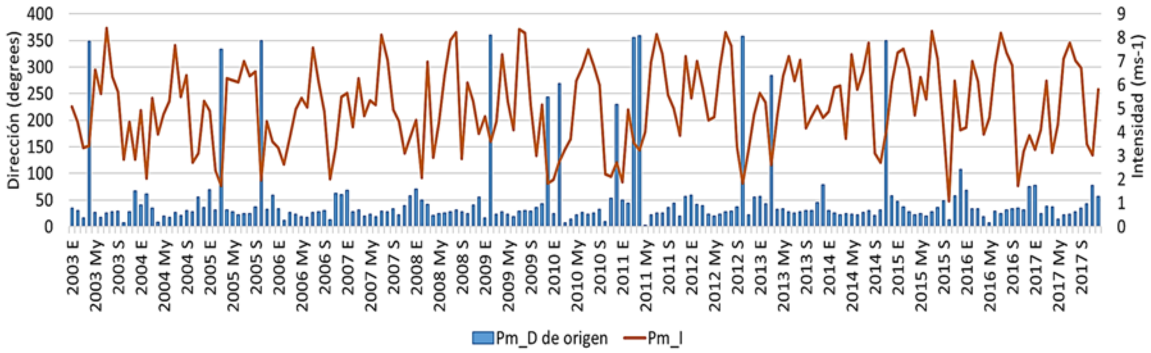


Fig. 3. Wind Intensity (Pm_I, orange line) and Wind Direction (Pm_D de origen, orange line) averaged time series of 8 data points. Monthly average. Period 2003-2017.

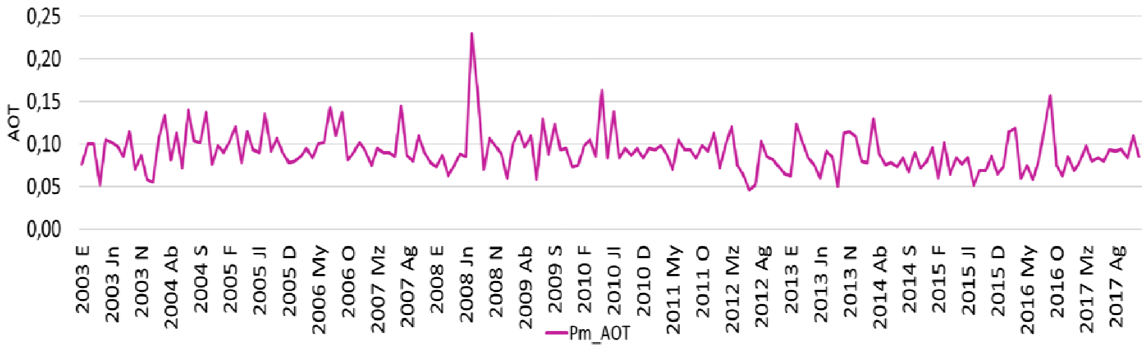


Fig. 4. Aerosol Optical Tickness (Pm_AOT) averaged time series of 6 data points. Monthly average. Period 2003-2017.

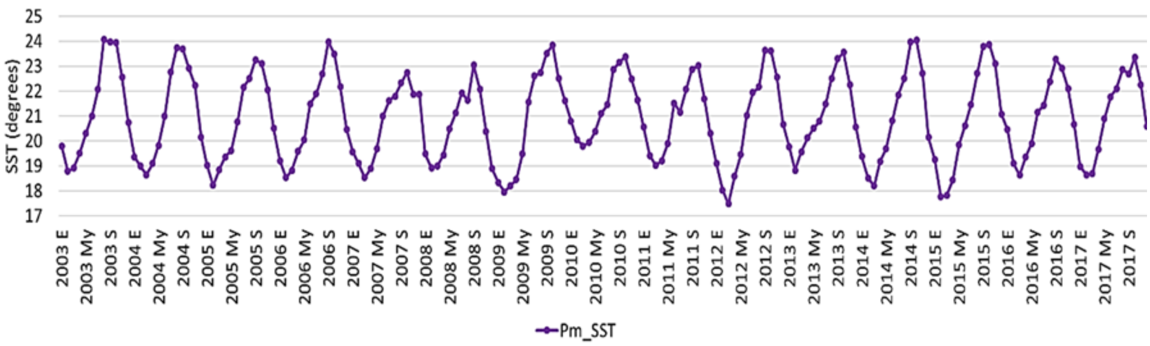


Fig. 5. Sea Surface Temperature (Pm_SST) averaged time series of 6 data points. Monthly average. Period 2003-2017.

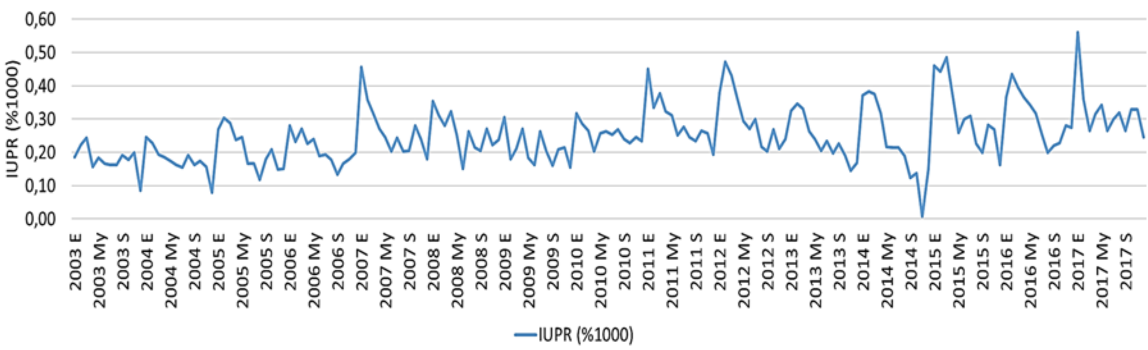


Fig. 6. Emergency admissions due to breathing pathologies (IUPR). Monthly accumulations. Period 2003-2017.

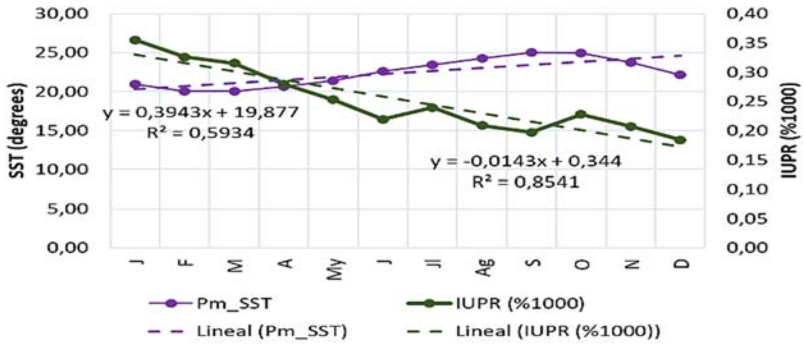


Fig. 7. Accumulated monthly averages of SST (·C, purple line) vs IUPR (% 1000, green line) and corresponding trendlines. Period 2003-2017.

As for the IUPR, data varies between 0.36 ‰ cases in January and a minimum of 0.18 ‰ in December.

The Pearson correlation coefficient between the monthly accumulated values of SST and IUPR (% 1000) for the 2003-2017 period is -0.661 ($p<0.001$).

Fig. 8 shows two of the climatic variables plot against each other, Wind Intensity (WI) vs the Aerosol Optical Thickness (AOT) both in monthly averages and for the full study-period. The correlation coefficient of this pair is the only one that, analyzed in (non-accumulated) monthly averages gives a significant statistical correlation, with Pearson's coefficient being 0.256 ($p<0.001$).

Fig. 9 shows the time series of AOT vs IUPR for monthly averages of the trimester March-April-May, for every year from 2003 to 2017. Visually, it is easy to identify that for most years, the AOT and IUPR present the same trends, in some cases (2006 and 2013) it is almost parallel. It is noticeable the year 2009 for the opposite event; the graphs are mirrored.

However, even though there is visually a clear relationship between the amount of aerosol in the atmosphere and the number of recorded respiratory medical admissions, the correlation coefficient is not statistically significant (0.083 (sig. (bilateral)=0.262)). Ongoing investigations are taking place in this respect.

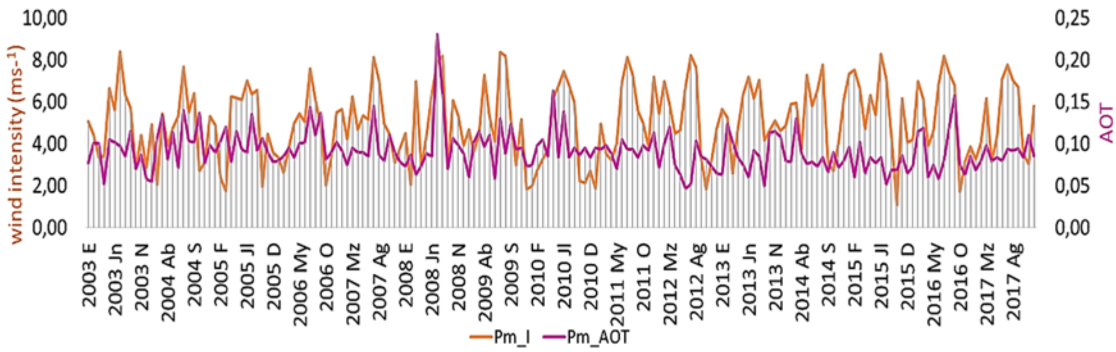


Fig. 8. Wind Intensity (ms-[1], orange line) vs AOT (pink line). Monthly average. Period 2003-2017.

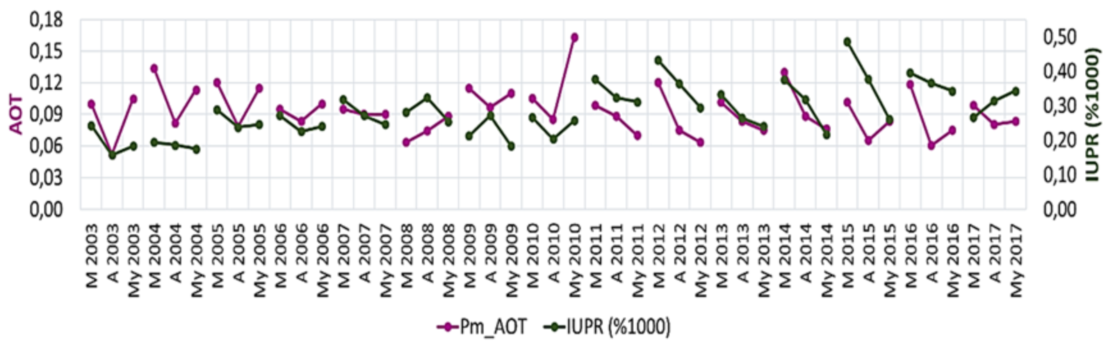


Fig. 9. AOT (adimentional, pink line) vs IUPR (% 1000, green line). Monthly averages for March, April and May. Period.

4. Conclusions

The results presented in this study show significant correlations between some climatic variables and the respiratory health status of the grancanarians. When looking into the correlation among the predicting variables (climatic variables), there is significant correlation between Wind Intensity and the Aerosol thickness. This coupled meteorological phenomena may be expected considering that the stronger the wind blows over the Sahara, the concentration of dust resuspended into the atmosphere increases as well as the stronger the wind, the further it is transported across the ocean and through the archipelago.

Analyzing the data in monthly averages for the 14 years period, the aerosols do no show significant correlation with any other climatic variable. In future studies the plan is to use wind data in a shorter temporal resolution (daily). Considering that the CALIMA events lasts in average 3 days at the time, using monthly averages might mask the data preventing, for instance, the expected correlation between wind direction and aerosol concentrations, to be significant.

Same applies to the correlation between Wind Direction and clinical data. It was expected to find a clear correlation between these two; aerosols reach the archipelago when the wind blows westwards or northwards over the Sahara. No correlation is shown in our study, so the idea is to repeat the correlation analysis using daily data.

On the other hand, it is also important to consider that the AOT data provided by MODIS-Aqua represents all solid and liquid particle in the full atmospheric column. Another option to depurate the data for correlation analysis would be to apply an algorithm to calculate the particles of a size between 2.5 μm and 10 μm, being this range the usually analyzed in human health related studies.

However, even though correlation analysis applied to the full length of the period do not show significance, when looking into the plot for shorter periods (i.e. trimester March-April-May) there seems to be a visual relationship between AOT and IUPR, showing a possible seasonal influence in the relationship of climatic and clinical variables. Further studies are carried out at the time this article is released.

When accumulating the data sets monthly for the full 14-year period, the variables SST and IUPR present significant correlation, however, it is important to note that this correlation does not necessary imply causality. The correlation between these two variables is slightly negative, which is expected considering that respiratory conditions tend to increase during winter, when the temperature drops, and decrease during the summertime.

For future studies, in order to get a better understanding of the influence of climatic factors like Wind Direction, it is advisable to use shorter temporal resolution, preferably daily time steps. In the case of Aerosols, considering working with specific particle sizes in advisable.

The final aim of this project is to help the Canary Islands health system be prepared for the changes that climate change might force in the number and nature of emergency admissions in the island of Gran Canaria. Considering the correlations found in this study, the well-known adverse impact of CALIMA in human health and the potential increase of CALIMA episodes due to climate change in the Macaronesian Region, it is essential to keep investigating the relationship between climatic and clinical parameters in the Canary Islands.

Acknowledgements

Use the Acknowledgements section if it is necessary. This work has been supported by the PASTOR (ProID2017010072) project, funded by the Department of Economy, Industry, Commerce and Knowledge of the Canary Islands Government, and the European Fondo Europeo de Desarrollo Regional (FEDER).

This study would not have been possible without the collaboration of the Admissions Team of the Gran Canarian Doctor Negrín University Hospital.

References

[1]. A. M. Petersen, E. M. Vincent, A. L. Westerling, Discrepancy in scientific authority and media visibility of climate change scientists and contrarians, *Nat. Commun.*, Vol. 10, Issue 1, 2019, pp. 1–14.

[2]. R. Haunschild, L. Bornmann, W. Marx, Climate Change Research in View of Bibliometrics, *PLoS One*, Vol. 11, Issue 7, 2016, p. e0160393.

[3]. G. Luber, M. McGeehin, Climate change and extreme heat events, *Am. J. Prev. Med.*, Vol. 35, Issue 5, 2008, pp. 429–435.

[4]. C. Rosenzweig, A. Iglesias, X.-B. Yang, P. R. Epstein, E. Chivian, Climate change and extreme weather events; implications for food production, plant diseases, and pests, *Glob. Chang. Hum. Heal.*, Vol. 2, Issue 2, 2001, pp. 90–104.

[5]. R. K. Pachauri, *et al.*, Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change, *IPCC*, 2014.

[6]. P. R. Epstein, Climate and health, *Science,* Vol. 285, Issue 5426, 1999, pp. 347–348.

[7]. W. R. Keatinge, G. C. Donaldson, The impact of global warming on health and mortality, *South. Med. J.*, Vol. 97, Issue 11, 2004, pp. 1093–1099.

[8]. L. C. Lau, K. T. Lee, A. R. Mohamed, Global warming mitigation and renewable energy policy development from the Kyoto Protocol to the Copenhagen Accord—A comment, *Renew. Sustain. Energy Rev.*, Vol. 16, Issue 7, 2012, pp. 5280–5284.

[9]. M. Pelling, J. I. Uitto, Small island developing states: natural disaster vulnerability and global change, *Glob. Environ. Chang. Part B: Environ. Hazards*, Vol. 3, Issue 2, 2001, pp. 49–62.

[10]. C. Toth, G. Jóźków, Remote sensing platforms and sensors: A survey, *ISPRS J. Photogramm. Remote Sens.*, Vol. 115, 2016, pp. 22-36.

[11]. L. R. Beck, B. M. Lobitz, B. L. Wood, Remote sensing and human health: new sensors and new opportunities, *Emerg. Infect. Dis.*, Vol. 6, Issue 3, 2000, p. 217-227.

[12]. Caballero Leiva Isabel Maria, Marrero Betancort Nerea, Rodriguez Betancor José Juan, Rodriguez-Esparragón Dionisio, Marcello Ruiz Javier, Correlation between Weather Conditions and Respiratory Pathologies in Gran Canaria by the Use of Remote Sensing, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'2019)*, Tenerife (Canary Islands), 25-27 September 2019, pp. 177–181.

[13]. S. Mallone, M. Stafoggia, A. Faustini, G. Paolo Gobbi, A. Marconi, F. Forastiere, Saharan dust and associations between particulate matter and daily mortality in Rome, Italy, *Environ. Health Perspect.*, Vol. 119, Issue 10, 2011, pp. 1409–1414.

[14]. S. Z. Sajani, *et al.*, Saharan dust and daily mortality in Emilia-Romagna (Italy), *Occup. Environ. Med.*, Vol. 68, Issue 6, 2011, pp. 446–451.

[15]. I. Menéndez, *et al.*, Saharan dust and the impact on adult and elderly allergic patients: the effect of threshold values in the northern sector of Gran Canaria, Spain, *Int. J. Environ. Health Res.*, Vol. 27, Issue 2, 2017, pp. 144–160.

[16]. M. Sanuy, E. Duo, W. S. Jäger, P. Ciavola, J. A. Jiménez, Linking source with consequences of coastal storm impacts for climate change and risk reduction scenarios for Mediterranean sandy beaches, *Nat. Hazards Earth Syst. Sci.*, Vol. 18, Issue 7, 2018, pp. 1825–1847.

[17]. L. Perez, *et al.*, Coarse Particles From Saharan Dust and Daily Mortality, *Epidemiology*, Vol. 19, Issue 6, 2019, pp. 800–807.

[18]. P. Kassomenos, A. Analitis, K. Katsouyanni, E. Kougea, E. Samoli, Does the presence of desert dust modify the effect of PM10 on mortality in Athens, Greece?, *Sci. Total Environ.*, Vol. 409, Issue 11, 2011, pp. 2049–2054.

[19]. J. M. Prospero, O. L. Mayol-Bracero, Understanding the Transport and Impact of African Dust on the Caribbean Basin, *Bull. Am. Meteorol. Soc.*, Vol. 94, Issue 9, Feb. 2013, pp. 1329–1337.

[20]. N. J. Middleton, A. S. Goudie, Saharan dust storms: nature and consequences, *Earth-Science Rev.*, Vol. 56, 2001, pp. 179–204.

[21]. Grupo de Procesado de Imágenes y Teledetección ULPGC, Servicios de estudios de la evolución climática de las caracteristicas del viento en la región Canarias-Souss Masaa Dràa, con imagen satélite en el marco del proyecto CLIMATIQUE, 2014.

[22]. J. Arístegui, P. Sangrá, S. Hernández-León, M. Cantón, A. Hernández-Guerra, J. L. Kerling, Island-induced eddies in the Canary Islands, *Deep Sea Res. Part I Oceanogr. Res. Pap.*, Vol. 41, Issue 10, 1994, pp. 1509–1525.

[23]. J. Carrillo, J. C. Guerra, E. Cuevas, J. Barrancos, Characterization of the Marine Boundary Layer and the Trade-Wind Inversion over the Sub-tropical North Atlantic, *Boundary-Layer Meteorol.*, Vol. 158, Issue 2, 2016, pp. 311–330.

[24]. L. Afonso, F. M. Galán, J. A. R. Martín, Geografía de Canarias, *Editorial Interinsular Canaria*, Vol. 1, 1984.

[25]. I. G. Nacional, Datos geográficos y toponimia, 2019. [Online]. https://www.ign.es/web/ign/portal/ane-datos-geograficos/-/datos-geograficos/. [Accessed: 12-Nov-2019].

[26]. A. J. McMichael, R. E. Woodruff, S. Hales, Climate change and human health: present and future risks, *Lancet*, Vol. 367, Issue 9513, 2006, pp. 859–869.

[27] E. Kalnay, The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc*., Vol. 77, 1996, pp. 437-470.

[28]. OBPG. NASA Goddard Space Flight Center, MODIS-Aqua Level-3 Binned Remote-Sensing Reflectance Data Version 2018, *2018 Reprocessing*, 2018. [Online]. https://oceancolor.gsfc.nasa.gov/l3/ [Accessed: 12-Nov-2019].

[29]. OBPG M. I. S. (MODIS) NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Aqua 11 µm Day/Night Sea Surface Temperature Data, *2018 Reprocessing*, 2018.

[30]. Dirección Gerencia del Hospital Universitario de Gran Canaria Dr. Negrín, Hospital Universitario de Gran Canaria Doctor Negrín, Memoria 08, 2009. [Online]. Available: https://www3.gobiernodecanarias.org/sanidad/scs/scs/as/gc/29/memorias/08/area_influencia.html. [Accessed: 12-Nov-2019]

Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 87-93

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

F-center Luminescence and Oxygen Gas Sensing Properties of AlN Nanoparticles

***Baiba BERZINA, Laima TRINKLER, Valdis KORSAKS, Rihards RUSKA, Guna KRIEKE and Anatolijs SARAKOVSKIS**
Institute of Solid State Physics, University of Latvia, 8 Kengaraga Str., LV-1063, Riga, Latvia
Tel.: +371-26157869, fax: +371-67132778
E-mail: baiber@latnet.lv

Received: 30 August 2019 /Accepted: 25 October 2019 /Published: 30 November 2019

Abstract: Luminescence caused by native defects in AlN nanoparticles (NP) with average diameter of 60 nm were studied using spectral characterization of the material. Photoluminescence and its excitation spectra were measured within a wide temperature interval between 300 – 8 K, and spectral range including ultraviolet (UV) and visible light. Luminescence properties were studied when a sample was surrounded by different media (vacuum and several gases such as oxygen, nitrogen, argon, air, oxygen-nitrogen mixtures). A blue luminescence (BL) consisting of two main sub-bands at 390 nm and 410 nm for AlN NP was observed. Two spectral regions around 315/340 nm and 260 nm were distinguished, including defect-induced absorption resulting in excitation of the BL. It was found that there are two different defect-induced luminescence mechanisms causing the BL. One of them is an intra-center mechanism and the other one is recombination luminescence mechanism, and their efficiency is highly dependent on spectral region of exciting light. It was found that oxygen gas which is surrounding the sample is reducing intensity of the BL. The oxygen gas sensing properties of AlN NP were studied allowing assumption that material is prospective for the development of oxygen gas optical sensors. As a result of this research the native defects responsible for the BL were related to nitrogen vacancy type defects known as the so-called F-centers. These defects can be located both on the surface and in bulk of the material, and the surface defects whose role is predominant in nanoparticles are responsible for oxygen sensing properties of AlN NP.

Keywords: Aluminum nitride, nanoparticles, luminescence, F-centers, oxygen gas sensing.

1. Introduction

At present elaboration and investigation of various luminescent nanomaterials is with great importance due to their small size and new excellent properties, which are decisive for the development of smart technologies and other applications. Large number of prospective material applications is based on their optical properties. Among these materials the wide band gap semiconductors are of special interest due to their absorption/luminescence spectra are covering a wide spectral range including visible and far ultraviolet (UV) light.

AlN is one of prospective wide band gap (E_g>6 eV [1-3]) materials with wurtzit crystalline structure [4] being accessible in forms of the bulk materials such as single crystals and ceramics, as well as in different types of nanostructures including nanopowders, nanorods, nanotips and others. Optical properties of the material are highly dependent on intrinsic defect types and the so-called point defects are the most important among them [5, 6].There is a large number of various defect types characteristic for the wide band gap materials including the so-called native defects which are inherent to the material without its special treating. The defects caused by the elements of the host material are the most typical among them. For AlN the most common are nitrogen and aluminum vacancies (v_N and v_{Al}, correspondingly), interstitials (i_N and i_{Al}) together with some native dopants such as oxygen and carbon atoms substituting for nitrogen (O_N, C_N) [7 and Refs. therein] and others. All these defects also can form different complexes.

The aim of the present study is determination of luminescent native defects inherent for AlN nanopowder together with revealing of luminescence processes and mechanisms caused by these defects as well as indication to some possible applications of this material based on its luminescence properties.

2. Material and Experimental

AlN nanopowder (NP) consisting of polycrystalline grains with average size of 60 nm was synthesized in Institute of Inorganic Chemistry, Riga Technical University by using plasma-chemical synthesis [8]. Raw aluminum powder (99.4 %) with particle size in the range of 20-40 μm was injected and evaporated into nitrogen (99.9 %) high-frequency plasma flow. In order to promote formation of AlN and to reduce particle growth ammonia was introduced additionally in high temperature flow containing aluminum vapors.

Photoluminescence (PL) and excitation (PLE) spectra of AlN NP were measured using a home-made setup, consisting of the following parts: *i)* a sample holder mounted inside of the sample chamber of a closed cycle refrigerator (CCS-100/204, Janis Research Corporation), allowing to perform spectral measurements at different fixed temperatures inside the interval 8 – 300 K, when sample is in vacuum, besides, the sample chamber can be filled with various gasses allowing measurements at room temperature; *ii)* a light source – a deuterium lamp (LD-400) for luminescence excitation and *iii)* a luminescence recording system containing the grating monochromator (Andor Schamrock SR-303i-B) together with the photomultiplier (Hamamatsu H7468-03) or CCD camera. All necessary spectral corrections related to the measurements are taken into account.

For spectral measurements AlN NP pressed in small tablets with 13 mm diameter and 1-2 mm thickness were used.

3. Results and Discussion

3.1. Native Defect Caused Luminescence of AlN Nanoparticles

Photoluminescence and its excitation spectra of AlN NP were measured for samples surrounded by vacuum (~10^{-5} mbar) at different fixed temperatures between room temperature (RT) and 8 K. Luminescence kinetics of PL also was studied at RT. Besides, the PL characteristics were examined when samples were surrounded with different gases including oxygen, nitrogen, argon, air and oxygen/nitrogen mixtures.

3.1.1. Photoluminescence Characteristics

Photoluminescence and its excitation spectra of AlN NP measured at RT when sample is in vacuum are depicted on Fig. 1, a and b, correspondingly.

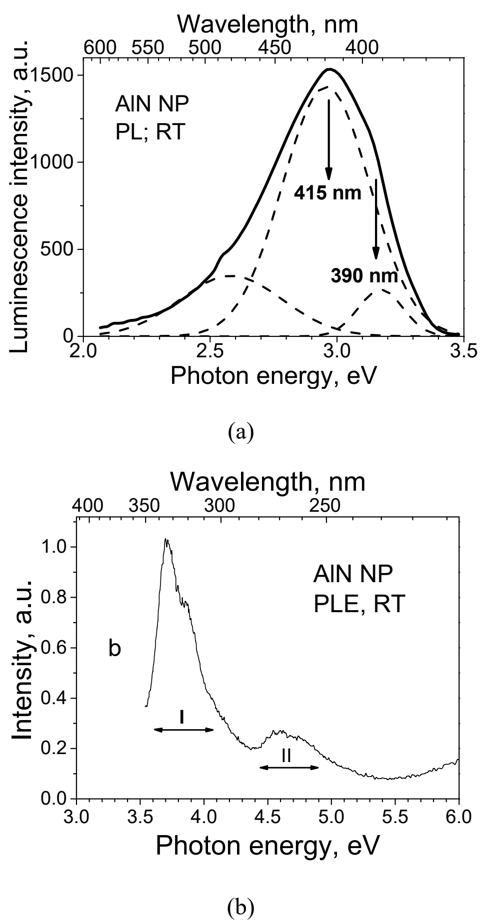


Fig. 1. AlN NP, RT. (a) PL spectrum at 315 nm or 260 nm excitation divided in three Gaussian curves (dash lines), (b) PLE spectrum of BL, where two spectral intervals are marked (I and II).

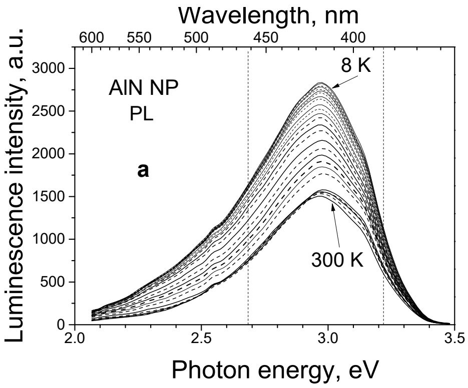
As it is seen from the Fig. 1, a, the PL spectrum forms a wide and complex band, which can be divided at least into three Gaussians depicted with dash lines. The main part of the PL spectrum includes the sub-bands peaking at 390 nm and 410 nm, which are situated within the near-UV and blue spectral regions forming together the blue luminescence (BL). Each of these BL sub-bands can be excited with light from a wide UV spectral range also containing several sub-bands (Fig. 1, b). In the long wave part of the PLE spectrum there are two main and highly overlapping bands at 340 nm and 315 nm, whereas, at the shorter wavelengths around 260 nm another wide and complex excitation band appears. Accordingly to this structure of the PLE spectrum, two different spectral intervals are marked as I and II on Fig. 1. b. Irradiation of material with light from the spectral region around 200 nm also results in appearance of the BL, which can be related to the exciton processes causing this luminescence. Excitation intervals I and II, with lower energies than that of exciton, are related to absorption of native defects. It is found that at RT the shapes of the BL spectra excited with light from the spectral intervals I and II are coincident.

PL spectra of ALN NP were measured at different fixed temperatures inside the range of 300 K – 8K.

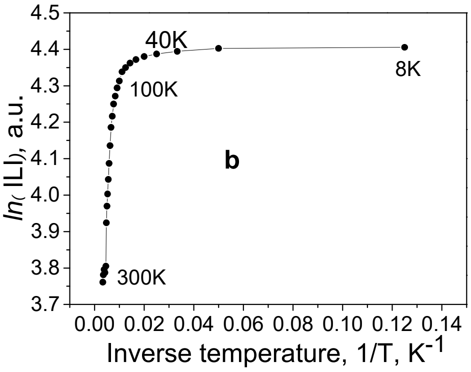
A set of the PL spectra, measured at different fixed temperatures under excitation of AlN NP with 315 nm light from the first spectral interval (Fig. 1, b) is shown in Fig. 2, a. It is seen that at various temperatures the shape of the luminescence band does not change essentially, whereas luminescence intensity is growing with decrease of temperature. These PL spectra measured at different temperatures of sample were used for building up Arrhenius plot of the integral luminescence intensity (ILI) on reverse temperature, which is seen on Fig. 2, b. For determination of the ILI the central part of the PL spectra framed by dotted straight lines depicted on Fig. 2, a, which includes the PL sub-bands at 290 nm and 410 nm, was used. As it is seen from the Fig. 2, b, there are two temperature intervals between 300 K – 100 K and 40 K to 8 K, where the graph is presented by a straight line. Inside of the high temperature interval 300 K - 100 K decrease of temperature results in increase of ILI and an observed linearity of the graph is characteristic for the simple exponential processes. Inside the low temperature interval from 40 K to 8 K the process is characterized by practically horizontal straight line, which demonstrates independence of the BL intensity on temperature. In general, such type processes described above are characteristic for the intra-center luminescence [9].

Dependence of the PL spectra of AlN NP on temperature under 260 nm excitation related to the II spectral interval of excitation (Fig. 1, b) was also investigated. In this case, besides the BL, which is observable at RT, another wide and complex luminescence band around 500 nm appears at temperatures below 200 K. Decrease of the sample temperature results in intensive increase of 500 nm luminescence intensity, which is highly exceeding that of the BL intensity thus modifying its spectral shape. The excitation spectrum of the 500 nm luminescence differs from that of the BL thus allowing assumption that both the BL and the luminescence at 500 nm are caused by various defect types.

Kinetics of the BL under excitation at 215 nm was studied at RT. It was found that after ceasing of exciting light the PL pulse is presenting an exponential decay, which can be characterized with short decay time in order of 10^{-7} s.



(a)



(a)

Fig. 2. AlN NP. (a) PL spectra at different temperatures measured within the interval between 300 K and 8 K and 315 nm excitation, (b) Arrhenius plot of the dependence of BL integral intensity (ILI) on reverse temperature. The used luminescence spectral interval is depicted by dotted lines.

During investigation of the PL spectra of AlN NP it was observed that the highest intensity of the BL appears when the sample is in vacuum, whereas, ambient air surrounding the sample decreases the BL intensity. A special study of this phenomenon was performed.

The BL spectra of AlN NP were measured at RT, when the sample is surrounded not only by vacuum, but also by different gases (Fig. 3). It was found that the BL of AlN NP is sensitive to surrounding oxygen gas, which is quenching the luminescence intensity (Fig. 3, curves 2 and 3). In the same time, other gases such as nitrogen and argon have not influence on the BL intensity, being concurrent with that observed when sample is in vacuum (Fig. 3, curve 1).

Sensitivity of BL to oxygen gas was tested and observed also in commercial AlN powder with average grain size exceeding 100 nm.

The observed phenomenon allows conclusion that the BL is mainly caused by the surface defects. When theses surface defects interact with surrounding oxygen a new type defects are created, which are different from those causing luminescence. Nevertheless, the BL is not completely quenched even when the sample is surrounded by pure oxygen gas

(Fig. 3, curve 3), thus allowing conclusion that in this case the observed BL is caused by the luminescent defects from in the bulk material of AlN grains. As it is seen from the Fig. 3, curves 1, 2 and 3, the shapes of all three PL spectra are coincident. It allows assumption, that the position of the luminescent defect being situated either on material surface or in its bulk does not affect significantly a structure of the luminescent defects.

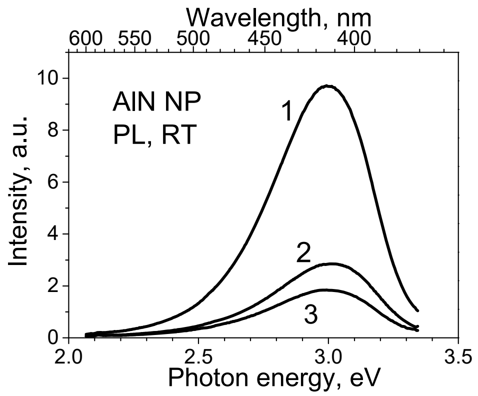


Fig. 3. AlN NP, RT. PL spectra under 315 nm excitation, when sample is in: 1 - vacuum (~10^{-5} mbar), nitrogen or argon gases; 2 – synthetic air; 3 – pure oxygen gas.

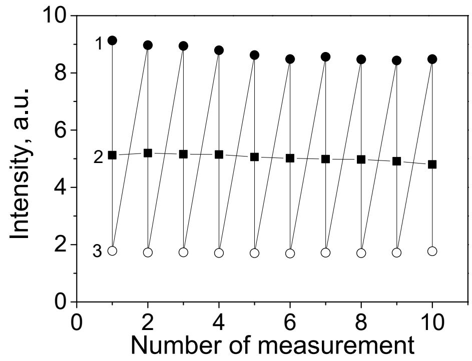
In summary, the results observed allow conclusion, that the defects which are causing the BL in AlN NP can be located in both the bulk of material and on its surface. In the case of nanomaterial the amount of surface defects exceed that from the bulk material due to extremely large specific area characteristic for the nano-grains.

3.1.2. Oxygen Sensing Properties of AlN NP

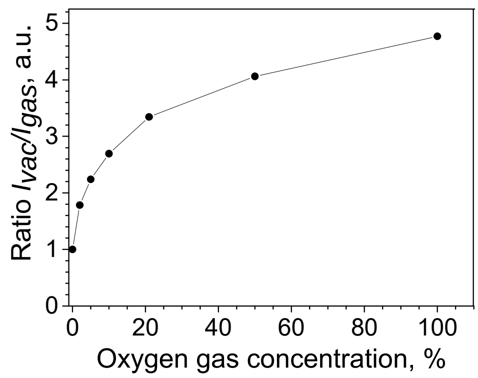
Oxygen sensing properties of AlN NP were examined. One of the most important characteristics describing oxygen gas sensing properties of AlN NP is repeatability of the response. This property is examined in the following way: the medium surrounding of the AlN NP sample was changed from vacuum to pure oxygen gas and backwards and this procedure was repeated ten times. In each of these states the luminescence spectrum was measured and its ILI was evaluated. The results are seen on the left side Fig. 4, where the ILI is depicted by the full circles (1) when sample is put in vacuum ~10^{-5} mbar or by the hollow circles (3) when the sample is surrounded by oxygen gas. The full squares (2) represent relation of the mentioned values. As it is seen from the Fig. 4 the repeatability of the results is good and AlN NP possesses qualities which are significant for oxygen gas sensors.

Dependence of the BL intensity on oxygen concentration in gas mixture surrounding the AlN NP material was studied. The commercial oxygen-nitrogen gas mixtures with different oxygen concentrations were used and the results are seen on Fig. 4, the right side. A nonlinear increase of the luminescent response, measured as ratio of ILI when sample is in vacuum and gas mixture, with increase of oxygen concentration is observed. The first point on this curve represents a case when the sample is in vacuum, but the last one – when the sample is surrounded with a pure oxygen gas. As it is seen from the picture the curve demonstrates a fast increase within a region of low oxygen concentrations below 10 %. Therefore, a more prospective use of the AlN NP as material for oxygen gas sensors could be related to the low oxygen concentrations (below 10 %) in gas mixtures. Nevertheless, the region of oxygen concentrations around 21 %, which is characteristic for ambient air, also seems to be considerable.

In summary, the observed spectral properties of AlN NP allow its ranking as the material useful for development of oxygen gas optical sensors.



(a)



(b)

Fig. 4. AlN NP; RT. (a) Repeatability of the response shown as luminescence intensity, measured when sample is in: vacuum ~10^{-5} mbar (1, full circles); in oxygen gas (3, hollow circles) and relation of 1/3 (2, full squares). (b) Sensitivity of the BL to oxygen gas concentrations in oxygen-nitrogen gas mixtures.

3.2. Luminescence Mechanisms

Analysis of the BL spectral characteristics of AlN NP discussed above allows revealing of several luminescence mechanisms, responsible for the BL, and each of them is highly dependent on conditions of luminescence excitation. For illustration of the processes, which are representing different luminescence mechanisms a scheme of AlN energy band structure and energy level states of two different defects (A and D) is depicted on Fig. 5.

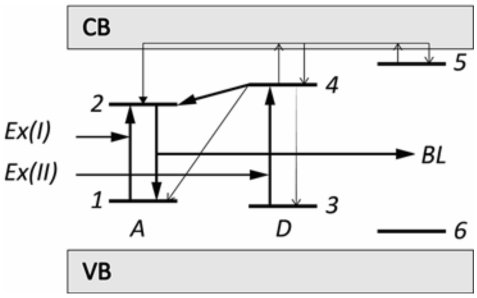


Fig. 5. Schematic representation of an energy band diagram of AlN (CB - conduction band and VB - valence band); energy levels of two different defects A and D, consisting of the ground states (1 and 3) and excited states (2 and 4), correspondingly, as well as trapping states for electrons (5) and holes (6). Bold vertical and oblique arrows are representing optical transitions between various defect states, but horizontal arrows are depicting the defect excitation (*Ex(I),* and *Ex(II))* and emission (*BL*).

It is possible suggest that for AlN NP there are two main native defect-induced mechanisms causing the BL. One of them is an intra-center mechanism, when the light absorption and emission occur inside one and the same defect. This mechanism is realized when the exciting UV light corresponds to the spectral interval I shown on Fig. 1, b. Existence of this mechanism is affirmed by the following facts. *i)*Position of the maxima of the luminescence bands and excitation bands from spectral interval I is very close and difference between them (the so-called Stock shift) does not exceed 1 eV (Fig. 1, a, and b). *ii)* Dependence of BL intensity on temperature within a region from 300 K to 8 K discussed above also affirms presence of intra-center processes. *iii)* The decay time of the BL in order of 10^{-7} s is characteristic for the intra-center processes. This intra-center luminescence mechanism is depicted on Fig. 5, realizing for the defect A, when absorption of light quantum (*Ex(I)*) induces optical transition 1→2 between the ground and excited energy states (1 and 2, respectively) and subsequent BL emission caused by transition 2→1. All processes occur inside one and the same defect A.

When the exciting light is corresponding to the spectral interval II (see Fig. 1, b), then the BL most credibly is caused by recombination processes. In this case there are two different but mutually interacting defect types, depicted as A and D on Fig. 5. One of them (D) is responsible for light absorption from the spectral region II (*Ex(II)*), but the other one (A) - for light emission (*BL*). Besides, in this case the defect A is the same as that one originating the intra-center luminescence, because the shapes of corresponding luminescence spectra are coincident at RT. Presence of luminescence recombination mechanism is confirmed by the following considerations. *i)* From the Fig. 1, a, and b it is seen that the positions of the BL emission bands and excitation band from the spectral interval II are considerably separated demonstrating a large value of the Stock shift, that is not typical for the intra-center processes. *ii)* In the same time, it is known that irradiation of bulk AlN with light from the II spectral interval results in realization of different recombination processes, including recombination luminescence of the so-called oxygen-related defects [10-14] and energy storage in the material, which can be released through thermally or optically stimulated processes [15-17].These observations together allow consideration that the recombination mechanism is causing the BL observed in AlN NP at irradiation with UV light around 260 nm from the spectral interval II depicted on Fig. 1, b.

Beside these two native defect-induced BL mechanisms an energy transfer mechanism from excitons to luminescent defect followed by BL emission is also present when the sample is excited with light from the far UV spectral region around 200 nm (see Fig. 1, b, a).

3.3. Luminescent Native Defects in AlN NP

Results of the present investigation together with an analysis of appropriate data from the scientific literature allow determination of defect types which are responsible for the BL in AlN NP.

At first we can exclude the well-known oxygen-related defects (ORD) observed in the bulk AlN, which also emit light from the blue spectral region, because the PL and PLE spectra of ORD [7, 10-14] essentially differ from those observed in the present study. Nevertheless, the spectral region of exciting light around 260 nm is common for both the ORD and the defects causing the BL in AlN NP. Therefore, we cannot completely exclude presence of ORD, when AlN NP is excited with light around 260 nm, but influence of ORD is physically impossible when sample is excited with light from the first spectral region seen on Fig. 1, b.

The most possible defect types which could be responsible for the BL in AlN NP are nitrogen vacancy type defects. Already in the early work of Y.C. Lan et al. [18] the PL at 477 nm in AlN powder was tentatively attributed to v_N centers. At present there is a lot of theoretical works stating that v_N forms a deep donor level in a wurtzite AlN structure [19-24]. Such position of defect state is necessary condition allowing assumption, that not only the v_N type defect ground state but also its excited state are placed inside the wide band gap of AlN. Then energy difference between these states could be big enough to cover spectral interval where the PL and PLE spectra are observed in the present study (see energy scheme on Fig. 5).

As is known, the v_N can capture electrons due to compensation of electrical charge in the crystalline lattice of the material thus forming the so-called F-centers [25]. For AlN there are three types of the F-centers named F, F^+ and F^{2+} which are created when 3, 2 or 1 electrons are added to the v_N. Such variety of the F-center types can explain the complex structure of the PL spectra.

In summary, we assign the defect types which are originating the blue luminescence in AlN NP to nitrogen vacancy type defects or the so-called F-centers. These defects can be located both in a bulk of material or on its surface. In the case of nano particles the surface defects can be dominant. There is a valid argument speaking in favor of this assumption. The blue luminescence is observed also in hBN and its features are very close to those observed in AlN NP [26]. Unifying element for both hBN and AlN is the nitride sub-structure of host material, where the v_N centers are formed.

4. Summary

Spectral characteristics of AlN nanopowder were studied. It was found that in AlN NP a defect-induced blue luminescence appears forming a wide complex band peaking at 415 nm, consisting of several sub-bands where the main of them are at 390 and 410 nm. This BL can be excited with light from two defect-related absorption regions situated around 315/340 nm and 260 nm. Depending on the excitation regions, two luminescence mechanisms causing the BL in AlN NP were found. One of them is the intra-center mechanism, which occurs when sample is irradiated with light from the long wave spectral region at 315/340 nm. The other one is the recombination luminescence mechanism, which can be activated with light from the short wave spectral region around 260 nm.

Analysis of our experimental results together with those reported in the scientific literature allow assumption that the native defects which are responsible for the BL luminescence in AlN NP can be related to nitrogen vacancy type defects known as the F-centers.

It was found that the BL in AlN NP is sensitive to surrounding oxygen gas which is quenching luminescence intensity. These investigations allow conclusion that the BL is caused by both the surface defects and those from the bulk material. In nanomaterial role of the surface defects is predominant. Good gas sensing properties such as repeatability of the response and others were found.
Sensibility of the BL of AlN NP to surrounding oxygen gas is ranking material among those being prospective for a development of oxygen gas optical sensors.

Acknowledgements

This work has been sponsored from the Latvian Sciences Council Grant No. Izp-2019/1-0443.

References

[1]. J. Li, K. B. Nam, M. L. Nakamuri, J. Y. Lin, H. X. Jiang, P. Carrier, and S. -H. Wei, Band structure and fundamental optical transitions in wurtzite AlN, *Applied Physics Letter,* Vol. 83, Issue 25, 2003, pp. 5163-5165.

[2]. E. Silveira, J. A. Freitas, S. B. Schujman, L. J. Schowalter, AlN temperature dependence from its optical properties, *Journal of Crystal Growth,* Vol. 310, 2008, pp. 4007-4010.

[3]. M. Feneberg, R. A. R. Leute, B. N. Neuschl, K. Tonke, M. Bickermann, High-excitation and high-resolution photoluminescence spectra of bulk AlN, *Physical Rewiev B,* Vol. 82, 2010, p. 075208.

[4]. H. Ott, Das Gitter des Aluminium Nitrides (AlN), *Zeitshrift fur Physic,* Vol. 22, Issue 1, 1924, pp. 201-214.

[5]. W. D. Palmer, The Elsevier Encyclopedia, Comprehensive Semiconductor Science and Technology, Ed. by Bhattachraya et al., *Elsevier Publisher Amsterdam,* Vol. 4, 2011, pp. 390-447.

[6] J. Bourgoin, M. Lannoo, Point Defects in Semiconductors II, *Springer Series in Solid State Sciences,* Vol. 35, 1983.

[7]. T. Koppe, H. Hofsass, U. Vetter, Overview of band-edge and defect related luminescence in aluminum nitride, *Journal of Luminescence,* Vol. 178, 2016, pp. 267-281.

[8]. J. Grabis, I. Steins, A. Patmalnieks, B. Berzina, L. Trinkler, Preparation and processing of doped AlN nanopowders, *Estonian Journal of Engineering,* Vol. 15, 2009, pp. 266-274.

[9]. B. I. Stepanov, V. P. Gribkovskii, Theory of Luminescence, *Gordon and Breach Publishing group*, 1968.

[10]. R. A. Joungman, J. H. Harris, Luminescence studies of oxygen-related defects in aluminum nitride, *Journal of American Ceramics Society,* Vol. 73, Issue 11, 1990, pp. 3238-3246.

[11]. S. Shweizer, U. Rogulis, J. M. Spaeth, L. Trinkler, B. Berzina, Investigation of oxygen-related luminescence centers in AlN ceramics, *Physica Status Solidi (B),* Vol. 219, 2000, pp. 171-180.

[12]. B. Berzina, L. Trinkler, J. Sils, K. Atobe, Luminescence mechanisms of oxygen-related defects in AlN, *Radiation Effects and Defects in Solids,* Vol. 157, 2002, pp. 1089-1092.

[13]. B. Berzina, L. Trinkler, J. Sils, E. Palcevskis, Oxygen-related defects and energy accumulation in aluminum nitride ceramics, *Radiation Effects and effects in Solids,* Vol. 156, 2001, pp. 241-247.

[14]. B. Berzina, L. Trinkler, D. Jakimovica, V. Korsaks, J. Grabis, I Steins, E. Palcevskis, S. Bellucci, L.-C. Chen, S. Chattopadhhyay, K. Chen, Spectral characterization of bulk and nanostructured aluminum nitride, *Journal of Nanophotonics,* Vol. 3, Issue 1, 2009, pp. 031950-16.

[15]. L. Trinkler, A. J. J. Bos, A. J. M. Winkelman, P. Christensen, N. A. Larsen, B. Berzina, Thermally and optically stimulated luminescence of AlN

ceremics, *Radiation Protection Dosimetry*, Vol. 84, 1999, pp. 207-210.

[16]. L. Trinkler, B. Berzina, A. Auzina, M. Benabdesselam, P. Iacconi, UV light energy storage and thermoluminescence in AlN ceramics, *Physica Status Solidi (C),* Vol. 4, No. 3, 2007, pp. 1032-1035.

[17]. L. Trinkler, B. Berzina, Recombination luminescence in aluminum nitride ceramics, *Physica Status Solidi B,* Vol. 251, No 3, 2014, pp. 542-548.

[18]. Y. C. Lan, X. L. Chen, Y. G. Cao, Y. P. Xu, L. D. Xun, T. Xu, J. K. Liang, Low-temperature synthesis and photoluminescence of AlN, *Journal of Crystal Growth,* Vol. 207, 1999, pp. 247-250.

[19]. A. Fara, F. Bernardini, V. Fiorentini, Theoretical evidence for the semi-insulating character of AlN, *Journal of Applied Physics,* Vol. 85, 1999, pp. 2001-2003.

[20]. C. Stampfl, C. G. Van de Walle, Theoretical investigation of native defects, impurities and complexes in aluminum nitride, *Physical Review B,* Vol. 65, 2002, p. 155212.

[21]. J. M. Vail, D. K. Chevrier, R. Pandey, M. A. Blanco, The nitrogen vacancy in aluminum nitride, *Journal of Pysics: Condensed Matter,* Vol. 18, 2006, pp. 2125-2135.

[22]. Yu. F. Zhukovskii, N. Pugno, A. I. Popov, C. Balasubramanian, S. Bellucci, Influence of F centers on structural and electronic properties of AlN single-walled nanotubes, *Journal of Physics: Condensed Matter,* Vol. 19, 2007, pp. 1-18.

[23]. K. Laaksonen, M. G. Ganchenkova and R. M. Nieminen, Vacancies in wurtzite GaN and AlN, *Journal of Physics: Condensed Matter* Vol. 21, 2009, p. 015803.

[24]. V. A. Soltamov, I. V. Ilyin, A. A. Soltamova, D. O. Tolmachev, E. N. Mokhov, P. G. Baranov, Identification of the deep-level defects in AlN single crystals: EPR and TL studies, *Diamonds and Related Materials,* Vol. 20, 2011, pp. 1085-1089.

[25]. J. H. Schulman, A. M. Stoneham, Color Centers in Solids, London, *Pergamon Press,* 1962.

[26]. B. Berzina, V. Korsaks, L. Trinkler, A. Sarakovskis, J. Grube, S. Bellucci, Defect-induced blue luminescence of hexagonal boron nitride, *Diamonds and Related Materials,* Vol. 68, 2016, pp. 131-137.

Sensors & Transducers
Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Rapid Strain Differentiation of *E. coli*-inoculated Urine Using Olfactory-based Smart Sensors

[1] Aminat Adebiyi, [1] Sarath Swaminathan, [2] Nam Than, [1] Mohammed Abdi, [1] Amy N. Bowers, [1] Andrea Fasoli, [1] Alberto Mannari and [1] Luisa Bozano
[1] IBM Research, Almaden, 650 Harry Road, San Jose, CA USA, 95120
[2] Department of Biomedical Engineering, San Jose State University, San Jose, CA USA
Tel.: + 14089271579
E-mail: aminat.adebiyi@ibm.com

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: Urinary tract infections (UTIs) have a large effect on the public healthcare system. Millions are affected each year, and are likely to have the infection recur with increased resistance to antibiotics. This is partly due to the lack of a targeted healthcare regimen that requires pathogen identification. As traditional cultures can take several days, there is a need for a real-time diagnostic tool for identifying uropathogens. In our study, the use of rapid olfactory-based smart sensors for the strain identification was explored by taking measurements of in-vitro assays involving the inoculation of two strains of *E. coli* (K12 and Uropathogenic *E. coli*) in urine. Using a Support Vector Machine classifier to train on two previous data collections yielded 95.83 % accuracy when tested on the third subsequent collection. These results demonstrate the promise of smart sensing in differentiating close metabolomic signatures of *E. coli* sand are expected to improve with more extensive data collection, assay development and sensor augmentation.

Keywords: Smart sensors, Electronic nose, *E. coli,* Urinary tract infection, Precision healthcare, Internet of things.

1. Introduction

With over 3 million cases a year in the United States alone, Urinary Tract Infections (UTIs) are the most prevalent bacterial infection and poses a great risk of morbidity in vulnerable populations, including children, pregnant women and the elderly [1]. In 2011, the estimated economic burden was estimated at $2.8 billion in the U.S. and the cost is only expected to increase with increasing hospitalization [2].

Accurate diagnosis of UTIs is difficult at the point of care, as it depends on both the presence of symptoms and a positive urine culture, which takes a minimum of 24 hours or up to multiple days for results [3]. A rapid test would lead to the correct empirical course of treatment targeted to the specific bacterial strain - significantly improving treatment outcomes, reducing medical costs and the likelihood of antimicrobial resistance.

Olfactory-based sensors like electronic noses have been widely studied and applied across industries to associate volatile organic compound profiles with the emitting analyte [4]. In healthcare, metabolomic markers can be used for studying disease diagnostics and has motivated the study of volatile organic compounds that correspond to these biomarkers [5].

In this paper, we introduce the Electronic Volatile Analyzer (EVA), a rapid, gas sensing IoT platform under development at IBM Research, Almaden - designed to deliver classification results under two minutes. A main objective of this study was to demonstrate the utility of rapid gas sensing IoT platforms in differentiating K12 and Uropathogenic (UPEC) strains of *Escherichia coli* (*E. coli*) inoculated into urine at comparable cellular counts.

2. Methods

2.1. The IBM Electronic Volatile Analyzer™

The IBM Electronic Volatile Analyzer™ (EVA, Fig. 1) is an electronic nose under development at IBM Research Almaden. The design of the platform is modular: each gas sensor is mounted on its own sensor module, a Printed Circuit Board (PCB) of common design carrying an integrated microcontroller and the circuitry needed to operate the sensor. The sensor modules communicate via I2C protocol with a central hub (BeagleBone Black), a single-board computer that orchestrates the sensor modules and processes the multisensorial output. A set of six commercial Metal Oxide (MOX) gas sensors was used to collect the measurements described in the following sections. The selection of the sensor portfolio was based on the indications of their respective manufacturers regarding the nominal target gases of each sensor. Consequently, the as-selected sensor array was expected to respond to a variety of volatile molecules, including alcohols, hydrocarbons, ammonia, methane, as a well as a wide range of volatile organic compounds (VOC), with a lower limit of detection at parts per million (ppm) level. Each MOX sensor was operated using an individualized multi-step, periodic voltage profile applied to the heating element, which resulted in step-wise modulation of the device temperature (Figs. 1b(i) and (ii)). Modulation of the temperature of MOX sensing elements is a well-known technique that can be used to enhance and tune the dependence of the sensing element resistance to the surrounding environment (Fig. 2b) [7]. The heater voltage profiles applied to the EVA sensors were optimized to maximize the sensitivity of each sensor with respect to a subset of target VOC, as well as to improve response orthogonality. Although the duration and amplitude of the individual waveforms was adjusted independently for each sensor, all waveforms were synchronized to a period of 80 s to simplify the handling and processing of the sensor array outputs. The resistance of each MOX sensing element was monitored at fixed voltage at a rate of 10 Hz. During operation, a constant flow of 150 ± 10 sccm was established by means of a small pump, drawing air from the environment in a continuous fashion while also collecting the headspace of the sample of interest. The vapor-carrying flow was directed towards an enclosed chamber containing the six MOX gas sensors, for the exposure and detection to take place.

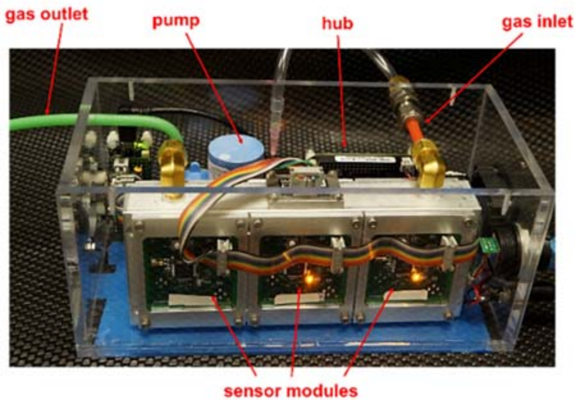


Fig. 1a. Electronic Volatile Analyzer (EVA) Sensor Array prototype with parts labeled.

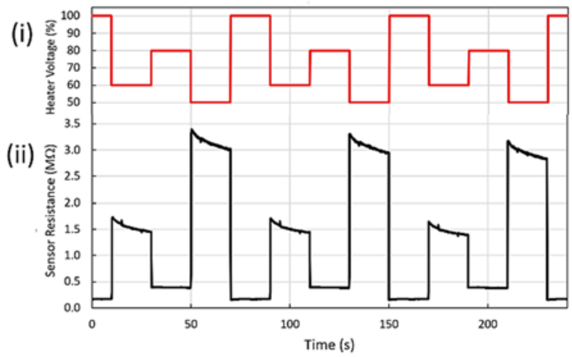


Fig. 1b. Example of a temperature profile modulation for MOX sensors for IBM EVA™: (i) Periodic waveform of heater voltage, expressed as a percentage of the maximum operating voltage recommended by the sensor manufacturer; (ii) Corresponding variations in MOX sensor resistance under constant environment.

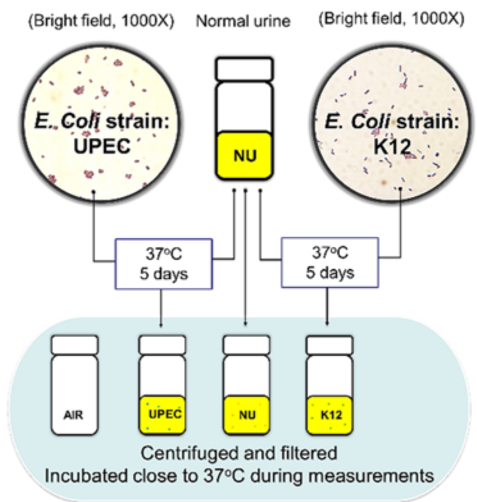


Fig. 2. Diagrammatic representation of assay development and conditions of data collection. Each strain of *E. coli* (K12 and UPEC) was inoculated separately into Normal Urine (NU) and incubated over five days, where it was filtered to trap volatiles. Each data collection involved four sample vials with Lab Air, NU, K12 and UPEC. Samples were incubated close to 37°C during measurements.

2.2. Assay Development and Sample Preparation

Two liquid UTI assays were developed by culturing *Escherichia coli* (Migula) Castellani and Chalmers (ATCC® 29425™) and *Escherichia coli* (Migula) Castellani and Chalmers (ATCC® 700928™) in Normal Human Urine (purchased from UTAK Laboratories). The Normal Human Urine (NU) was centrifuged and filtered with 0.2 µm filters prior to inoculation, to remove sediments and to aid in the homogeneity of samples. Each of the *E. coli* strains were then separately inoculated into NU at a fixed concentration, and incubated at 37 ºC over five days to guarantee a terminal concentration of 10^9 CFU/ml. After samples were plated, they were further filtered to remove cellular bodies of the bacteria, leaving only their emitted metabolomic volatiles behind [6].

Sample volumes of 5ml were dispensed in 20ml Wheaton Septa Top vials for measurement. The list of samples measured were filtered Uropathogenic *E. Coli* (UPEC) in NU, K12 *E. coli* in NU, Normal Urine, and Lab Air as reference.

2.3. Data Collection and Processing

Collection of data for the training and testing datasets were made at different times, with a minimum of eight hours gap in between. The training dataset was comprised of two experiments, and the test dataset was the third sequential acquisition. Headspace measurements of each of the four sample vials were made using the IBM EVA™. The sequence of collection from the sample vials was randomized over the course of an eight hour collection period, with an air flush between vials. Samples were incubated close to 37ºC during measurements, to aid in the evaporation of trapped volatiles. The response of the sensors was saved using Data Acquisition System (DAS) before feature extraction and labeling of the sensorial data were performed. Various feature extraction approaches were explored on sensor raw resistance curves to capture sensitivity and amplitude differences in the samples. Machine learning was performed mainly on the extracted feature set to identify the odor classes from the testing dataset (Testing Batch) by training on the two prior collections of the training set (Training Batches A and B), which is the main analytical model in view of the intended application. Given the limited number of total examples, traditional machine learning classifiers were selected, including Support Vector Machine (SVM), Logistic regression, Random Forest and Multilayer Perceptron. Implementation of these classifiers was from the Scikit-learn 0.19.1 package on Python [8]. The entire dataset was cross validated using three-fold validation with each batch serving as a fold – also referred to here as leave-one-batch-out. The parameters for cross validation are shown in Table 1.

Table 1. Cross validation parameters for selected traditional machine learning classifiers

Classifier Type	Tuning Parameters
Logistic Regression (LR)	‘l1’ and ‘l2’ loss; C values from 0.001 to 1.0; optimization algorithms – ‘newton-cg’, ‘libfgs’, ‘liblinear’, ‘sag’, ‘saga’; multiclass classifications using ‘one-vs-rest’ and ‘multinomial’
Support Vector Machine (SVM)	‘linear’ and ‘rbf’ kernels; C values from 0.001 to 1.0; ‘one-vs-one’ and ‘one-vs-rest’ for multiclass classification
Random Forest (RF)	estimators from 5 to 20; maximum depth of tree from 2 to 10 or None; max_features - ’auto’, ‘sqrt’, ‘log2’ and None; and criterion – ‘gini’ or ‘entropy’
Multi-layer Perceptron (MLP)	one to two hidden layers with size ranging from 10 to 25; sigmoid and relu activation functions; stochastic gradient descent and adam solvers; l2 penalty from 0.0001 to 0.01

3. Data Analysis

3.1. Data Visualization

Resultant data from the projected input set by sample label is shown in Fig. 3. The two main principle components of the dataset represented 98.7% of variation in the data, showing the dataset is separable and easily defined. Training Batch A showed more variation across labels compared to Training Batch B and Testing Batch. Based on the sequence of collection, measurement drift was experienced in Training Batch A but stabilized between Training Batch B and Testing Batch. This phenomenon could also represent the correlation of concentration of volatiles in samples between the batches.

3.2. Features

The raw resistance sensor output values are high dimensional and noisy and are not suitable to be directly used as input for classifiers. To transform this raw values into usable features and thereby to represent the data better for machine learning algorithms, we explored the following feature extraction approaches:

Max: Maximum response values for sensors corresponds to their degree of reaction to the gas. For each sensor, the resistance curve is equally divided into ten sections and extracted the mean resistance value for each sections.

Fractional Difference: Fractional difference is an another kind of maximum value feature, which is defined to correct additive errors in the raw resistance values. For each sensor, the maximum resistance value (Rmax) and minimum resistance value (Rmin) are taken. The fractional difference value for a sensor is defined as (Rmax – Rmin) / Rmin.

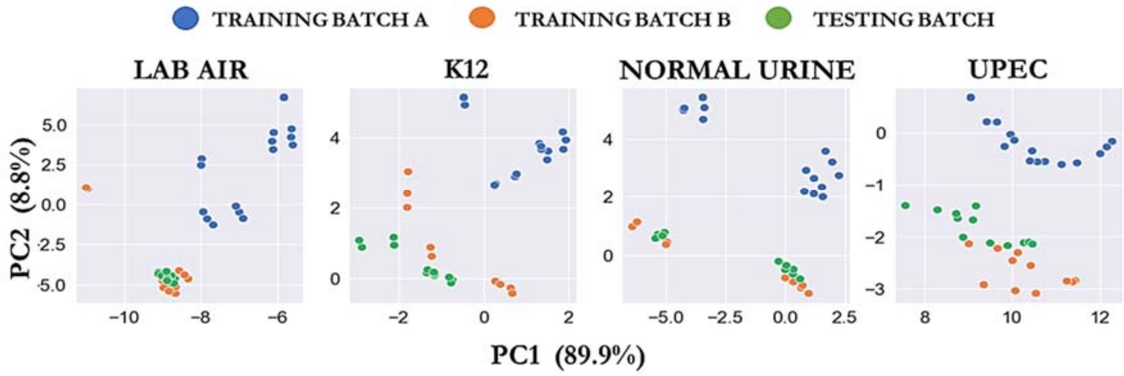


Fig. 3. Principal Component Analysis (PCA) visualization of each sample label (Lab Air, Normal Human Urine with K12 E.coli strain, Normal Human Urine and Normal Human Urine with UPEC *E. coli* strain) by data collection batch. The two axes represent Princpal Components (PCs) 1 and 2 which describes 98.7% percent of the variation in the data. Training Batches 1(blue) and 2 (orange) were collected in sequence before the terminal collection of the Testing Batch (green). Both Training Batches were used to develop the classifier for which the Testing Batch was tested on. Training Batch B and Testing Batch overlap closely, while Training Batch A is more distant with respect to the other batches.

First Derivative: To smoothen the resistance curve for each sensor, the curve is divided into ten equal sections and the mean is taken. The first derivative is then taken over this "smoothened" mean representation.

Mean: For each sensor, the resistance curve is equally divided into ten sections and the mean resistance value for each section is taken. This ten-dimensional mean representation of the raw data, smoothens the initial raw resistance curve of sensors.

PCA: This common statistical technique was used to reduce dimensionality of mean values to obtain a six-dimensional representation of the sensor data.

Effectiveness of these features was measured by finding the average accuracy of the models with leave-one-batch-out classification. Initially the best parameter for the classifiers were tuned for each of these features. Cross validation based on the parameters list in Table 1 was used for parameter selection. Leave-one-batch-out or three-fold validation was done by training on two batches and testing on the third batch. This was repeated where each of the three batches served as the Testing Batch at least once. Average test accuracy for all three batches of data is reported in Table 2.

Based on the accuracies from Table 2, mean and max features performed equally well while fractional difference fell short. And one interesting thing to note is when combined mean and max features accuracy reduced significantly. PCA features did slightly better than max and mean features. First derivative feature gave surprisingly great result and when combined with max features it improved further. As we can see from Table 2. *Max & First derivative* feature gave perfect result of accuracy 1.0 for SVM, Random Forest and Multi-layer Perceptron. And Random Forest gave slightly better result for other features compared to other classifiers.

Based on the tuning parameters the best parameters for each classifier is outlined in Table 3. Given the linearity of the data shown in Fig. 3, the cross validation parameters for classification show very high separability based on the high accuracy achieved for all feature types and models attempted, and is confirmed through the PCA visualization.

Table 2. Average leave-one-batch-out accuracy scores for selected classifiers and features

Features	LR	SVM	RF	MLP
Mean	0.91	0.96	0.91	0.95
Max	0.94	0.99	0.93	0.95
First Derivative	1.00	1.00	1.00	0.96
Fractional Difference	0.84	0.87	0.92	0.87
PCA	0.91	0.98	1.00	0.93
Mean & Max	0.85	0.85	0.90	0.92
Max & First Derivative	0.98	1.00	1.00	1.00
Average	**0.92**	**0.95**	**0.95**	**0.94**

Table 3. Best parameters for classifier type for training data

Classifier Type	Best Parameters
Logistic Regression	C=0.1, multi_class='multinomial', penalty='l2', solver='lbfgs'
Support Vector Machine	C=1.0, decision_function_shape='ovr', kernel='linear'
Random Forest	criterion='gini', max_depth=5, max_features=n_features, n_estimators= 10
Mult-layer Perceptron	activation='relu', alpha=0.01, hidden layer 1 = 20, hidden layer 2 = 10, solver= 'adam'

The detailed results of the performance of the classifiers on the Max and First derivative feature set for the three-fold validation with each batch serving as a fold is shown in Table 4. Table 4a shows the accuracy of using the indicated batch as a hold out while training on the remaining batches. Table 4b shows the F1 scores, which is the harmonic mean of the precision and recall, measuring the accuracy of the classifier on the dataset. In addition to the main sequential model for our application, the results for the remainder of the batches as hold-outs are shown.

Table 4. Normalized accuracy and F1 scores of leave-one-out batch by classifier type when trained on two other batches and tested on indicated sample set with Max and First Derivative as feature type.

(a) Normalized Accuracy Scores

Batch	LR	SVM	RF	MLP
Training Batch A	0.94	1.00	1.00	1.00
Training Batch B	1.00	1.00	1.00	1.00
Testing Batch	1.00	1.00	1.00	1.00
Average	**0.98**	**1.00**	**1.00**	**1.00**

(b) F1 scores

Batch	LR	SVM	RF	MLP
Training Batch A	0.93	1.00	1.00	1.00
Training Batch B	1.00	1.00	1.00	1.00
Testing Batch	1.00	1.00	1.00	1.00
Average	**0.98**	**1.00**	**1.00**	**1.00**

For the main sequential model the application is based on, which involves training on Training Batches A and B and testing on the last collection, all classifiers performed perfectly with an accuracy and f1 score of 1.0. While all classifiers performed well on all combination of leave-one-batch-out testing, Logistic regression did slightly worse for predicting Training Batch A.

All the models performed very well on average for all the feature types, with over 90 % accuracy. SVM and Random Forest, performed the best with a tied 95 % average accuracy. The feature combination of Max and First Derivative proved to be only slightly better that the First Derivative on its own, which performed the worst using Logistic Regression as a classifier, while the Max/First Derivative combination performed worst using Multi-Layer Perceptrons. Otherwise, both feature types achieved a 100 % classification accuracy with SVM and Random Forest models.

4. Conclusions

The results of the pilot experiments presented here highlight the promise of olfactory-based smart sensors in pathogen identification for bacteria of the same genus. Further studies, including extensive data collection, assay development and sensor augmentation are warranted to validate this device as a precision antibiotic matching tool.

References

[1]. A. Flores-Mireles et al., Urinary tract infections: epidemiology, mechanisms of infection and treatment options, *Nature Reviews Microbiology,* 13, 5, 2015, pp. 269-284.

[2]. J. E. Simmering et al., The increase in hospitalizations for urinary tract infections and the associated costs in the United States, 1998–2011, *Open Forum Infectious Diseases,* Vol. 4, No. 1, 2017.

[3]. B. Foxman, Epidemiology of urinary tract infections: incidence, morbidity, and economic costs, *The American journal of Medicine,* 113, 1, 2002, pp. 5-13.

[4]. A. Wilson and B. Manuela, Applications and advances in electronic-nose technologies, *Sensors,* 9, 7, 2009, pp. 5099-5148.

[5]. A. Amann and S. David eds., Volatile biomarkers: non-invasive diagnosis in physiology and medicine, *Elsevier Science Ltd*, 2013.

[6]. R. B. Caldwell and T. L. Claus, Simple and reliable method to precipitate proteins from bacterial culture supernatant, *Appl. Environ. Microbiol.,* 70. 1, 2004, pp. 610-612.

[7]. A. Schutze, T. Baur, M. Leidinger, W. Reimringer, R. Jung, T. Conrad, T. Sauerwald, Highly sensitive and selective VOC sensor systems based on semiconductor gas sensors: how to?, *Environments,* 4, 20, 2017.

[8]. F. Pedregosa, G. Varoquaux, et al., Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research*, 2011, 12, Oct, pp. 2825-2830.

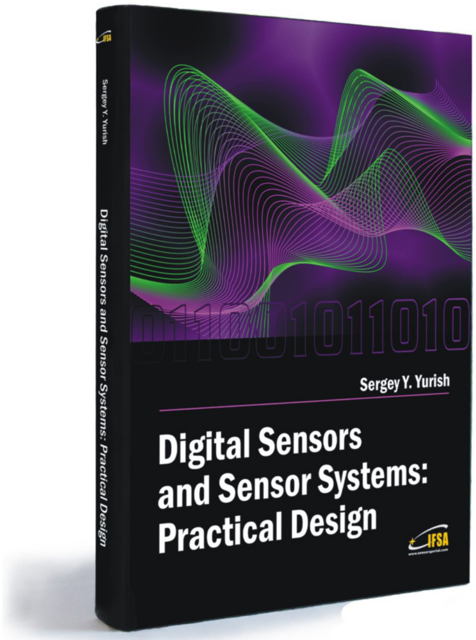
Theory:

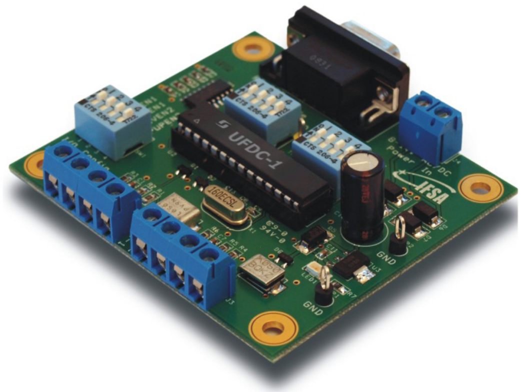
Digital Sensors and Sensor Systems: Practical Design

and

Practice:

Development Board EVAL UFDC-1/UFDC-1M-16





Buy book and Evaluation board together. **Save 30.00 EUR**

Development Board EVAL UFDC-1 / UFDC-1M-16

Full-featured development kit for the Universal Frequency-to-Digital Converters UFDC-1 and UFDC-1M-16. 2 channel, 16 measuring modes, high metrological performance, RS232/USB interface, master and slave communication modes. On-board frequency reference (quartz crystal oscillator). Operation from 8 to 14 V AC/DC. Development board software is included.

All existing frequency, period, duty-cycle, time interval, pulse-width modulated, pulse number and phase-shift output sensors and transducers can be directly connected to this 2-channel DAQ system. The user can connect TTL-compatible sensors' outputs to the Development Board, measure any output frequency-time parameters, and test out the sensor systems functions.

Applications:

- Digital sensors and sensor systems
- Smart sensors systems
- Data Acquisition for frequency-time parameters of electric signals
- Frequency counters
- Tachometers and tachometric systems
- Virtual instruments
- Educational process in sensors and measurements
- Remote laboratories and distance education

Order online:
http://www.sensorsportal.com/HTML/BOOKSTORE/Digital_Sensors_and_Board.htm

IFSA

International Frequency Sensor Association Publishing

www.sensorsportal.com

ISSN 1726- 5479

9 771726 547001