Application Of Mathematical Models And Digital Filters And Their Processors Of Spectral Analysis For Aromatic Compounds Gas In A Fluorescent Chemical

Al-Ibadi Zeyad\textsuperscript{a}, Muthana Alboedam\textsuperscript{b}, Ilya KATANOVC, Al-Zubaidi Sura\textsuperscript{d}

\textsuperscript{a}Faculty of Informatics, Department of Technical Cybernetics, Samara State University, Samara, RUSSIA, st. Molodogvardeyskaya, 151, Building 1.
\textsuperscript{b}Faculty of Informatics, Department of Technical Cybernetics, Samara State University, Samara, RUSSIA, st. Molodogvardeyskaya, 151, Building 1.
\textsuperscript{c}Faculty of Informatics, Department of Technical Cybernetics, Samara State University, Samara, RUSSIA, st. Molodogvardeyskaya, 151, Building 1.
\textsuperscript{d}Faculty of Biology, Samara State University, 443011, Russia, Samara, Building 22b, st. Academician Pavlova, 1.

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: In this work, different spectra processing methods are affected by the principal components (Aromatic compounds). Keeping up high-spatial objectives is progressively basic. Various methodologies are utilized to check fragrant compounds that anticipate choosing a specific technique that's amid research Center determinations. These techniques empower us to assess a particular degree of normal compounds, much the same as benzene, toluene and xylene, and so on. One notable part of all types of signal systems is the flexibility of adaptation. And spatial exactness isn't fundamental to get a range from an expansive number of fragrant compounds where more prominent characterization and statistical mean are more critical. Moreover, sufficiently low deviations of the expected values were achieved from the true values. The standard deviation, to determine the properties of fragrant compounds and compare them with normal compounds isn't thorough. A persistent baseline rectification was performed; after that, the rectified spectrum was normalized to their area and somewhat smoothed. The autofluorescence foundation was subtracted, for the pure range analysis, by utilizing scientific approaches: polynomial estimation (PolyFit) and (method Processors Gases Improved). The accuracy obtained is not extreme and can be increased by developing algorithms and selecting other parameters. It is also possible to increase the accuracy and reliability of this method by improving the quality of the training sample by eliminating the unwanted data that we have obtained, by increasing the sample size, and by studying more in detail the sample data to eliminate inaccuracies that arise during the transition between concentrations Gas.

Keywords: PolyFit, Polynomial, Processors Gases, Ultraviolet Spectroscopy, Sensors, Fluorescence.

1. Introduction

Aromatic compounds are broadly utilized in cutting edge industry. It is necessary to know their concentration in the air, particularly in industrial and populated ranges, due to their harmfulness indeed at parts per billion concentrations [1], these compounds such as benzene, toluene and xylene, cause extraordinary hurt to human health,(for example, anaemia, sterility and a few sorts of cancers), in addition to environmental dangers, despite the health issues caused by these aromatic compounds, we note that the production of these compounds increments universally [2].

In the last twenty years have used sensors the analytical regarding chemical compounds in the environment in all diverse branches of industry and medicine. Different sensing is developed and used in numerous medical-related and biochemical applications to the safety systems food-packaging sector [3].

To perform a dedicated look for optical-sensor materials and even produce them, one should be capable of computing the spectral properties in addition to functional characteristics of typically the sensor materials under adjustable and reproducible conditions [4].

Optic sensors based on fluorescent reagent-mediated principle usually be made up of sensor stuff and a sensor program. The sensor material includes indicator molecules immobilized over a polymer or other form of a matrix [5].

The sensor stage is closely related to the physical principle used to study the sensor material's interaction and the manufacturing technique. The reliable components of the fluorescence reading platform are the source of light to excite the fluorescence, in addition to the filtered fluorescent light isolated from excitation light and photodetector.
Yet, this method features certain limitations, such as the need to use numerous chemical sensory elements concurrently, making calibration difficult, which requires getting alternative ways to overcome these types of difficulties [6].

In this study, the algorithm was developed to solve the fluorescence spectrum analysis [7]. This examine's major idea was to use the particular information extracted from the spectral data and increase the accuracy of calculating the particular concentration of aromatic hydrocarbons in the air typically. This allows one to efficiently spectroscopy of the gas fluorescence, determine the concentration [8].

2. Material and Method

2.1. Experimental Setup

(Dibenzoylmethanato) boron difluoride was prepared totally to the method portrayed by Brown and Bladon [9]. The test checks of multicomponent blend concentrations; they were recorded using a (sensor) sensor fabric blended on a glass slide. The complex optical fiber test is associated with the back of the chip. The glass slide with the sensor fabric is found in a fixed cell in which clean air or a mixture of air is provided with aromatic hydrocarbon vapours at a certain concentration. The sensor is associated with ultraviolet (LED Nichia, λ = 375 nm), and detection with the Ocean Optics S2000 fiber spectrometer, and Filtration of detected radiation with LeeFilters 226 LEE U.V. a 200-millisecond pulse signal detection scheme and a 185 second period were used.

2.2 Specimens Preparation

The with sets of aromatic hydrocarbons given examples were prepared from the laboratory (Samara State University). Combinations of the concentrations, Barometers were Gas supply, for Naphthalene (0.2 μg / min), Benzene (44.0 μg / min), Toluene (10.1 μg / min), Xylene (2.4μg /min), and temperatures at which the IMP is maintained for Naphthalene (50° C), Benzene (110° C), Toluene (100 ° C), Xylene (90°C). Air purge flow, ml/min 200. Were introduced into the cell, and direct recording of spectral characteristics, altogether we performed our study for more than 1300 samples. Experimental studies of the samples were performed within 185 seconds from the minute of testing. A standardized recording of the spectra was carried out. It should be noted that it is vital to neglect values tests taken and all estimations at which the concentration of all gasses was zero from the study.

The information and storage of major components compatible with the spectral characteristics information were preprocessed with MATLAB (R2013a, Math Works).

2.3 (ProcessorsGases) Algorithms, Filtering Signals for Removing the Background of Autofluorescence

![Flowchart of data processing](Image)
The whole signal processing process can be divided into six stages. It all starts with the fact that with the help of the data logger, we get real “raw” spectra. In this work, aromatic gas sensor spectra are used; in the next stage, our attention was focused on these spectra groups. The cut spectra are still very noisy. In the third stage of data processing, we remove noise to make spectra more useful, and in the final stage, another non-informative parameter is deleted. After the background is successfully removed, the “pure” spectrum is obtained, consisting of peaks corresponding to the time, (or wavelengths) in which molecular vibrations occur in the test sample. A diagram of the algorithm in Figure 1.

A monochromatic light occurrence on a sample is characterized by phenomena such as reflection, scattering, or absorption.

In the common case, the scattering process competes with its absorption process. When radiation is absorbed, the particle moves to a less energizing electronic state. The return move to the ground state can be either lighted or went with by light radiation at a lower recurrence. This radiation is called luminosity (Luminescence). And through mathematical models for recording signals, the signal received from the sensor can be spoken to by a some included scientific models that comprise the spectrometer signal and the foundation of self-sparkle and clamber (noise of all sorts) [10].

Because the noise filtering nature in signals is diverse, we apply a median filter to remove impulse noise at the preliminary stage. Its main advantage is that it reacts with less quality to readings that stand out sharply against the background of neighbouring ones, which allows it to be widely used to eliminate abnormal values in data arrays [11].

Also, the median filter turns out to be more efficient than linear filters in cases when the differences in signal values are large compared to the variance of the Gaussian noise as it gives a lower value of the root mean square error of the output signal to the input, noise-free signal when compared with optimal linear filters [12].

In general, as the arithmetic means of two averages. The output signal \((y_k)\) of a sliding median filter with a width of \((2n + 1)\) for the current sample \(k\) is formed from the input time series \((\ldots, x_{k-1}, x_k, x_{k+1}, \ldots)\) following formula (Equation 1):

\[
y_k = M e(x_{k-n}, x_{k-n+1}, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_{k+n-1}, x_{k+n})
\]

(1)

Where \((x_1, x_2, \ldots, x_{2n+1}) = x_{n+1}, x_m\) elements of the variation series, i.e. ranked in ascending order of values \(x_m\):

\[
x_1 = \min(x_1, x_2, \ldots, x_{2n+1}) \leq x_2 \leq \ldots \leq x_{2n+1} = \max(x_1, x_2, \ldots, x_{2n+1})
\]

The width \((2n + 1)\) of the median filter is chosen to suppress a pulse with a width of \(n\) samples. The smoothing procedure is carried out using the Savitsky-Golay filter.

The Savitsky-Golay filtering method is better than simple averaging because it retains the spectrum's characteristic features, such as the peaks' height and width.

This approach's main advantage is that it tends to maintain local extrema, which are usually smoothed out by other averaging methods. The formula for Savitsky-Golay filter (Equation 2):

\[
S_i = \sum_{j=-n}^{n} c_j * S_{i+j} \quad (i = n + 1, N - n)
\]

(2)

where \(S_i\) - samples of the filtered signal; \(c_j\) - coefficients of Savitsky-Golay; \(n\) - is the half-width of the smoothing window. Savitsky-Golay coefficients are calculated by the formula (Equation 3):

\[
c_j = C_{i,j} \quad (j = 1, 2n + 1)
\]

(3)

where \(C = (A^T A)^{-1} A^T\)

\[
A_{i,j} = \begin{cases} 
(i-n-1)/i, & i \in [1; n] \cup [n + 2, 2n + 1], \; j \in [2, K + 1] \\
1 = n + 1, & j = 1
\end{cases}
\]

The width \((2n + 1)\) and the regression order \(k\) of the smoothing filter are chosen to not lead to significant distortion or loss of informative signal components [11]. The overall filtering scheme is shown in Figure 2.
Background removal is one of the most difficult tasks for a qualitative analysis of experimental results. A variety of techniques encompass either instrumental or software approaches to remove background fluorescence signals, software methods have become the standard for correcting the contribution of fluorescence to the background. They do not require modification of the system and have no restrictions on sample preparation. The mathematical methods, first and second derivative methods, frequency filtering, polynomial approximation, and wavelet transform are suggested as useful background removal methods in certain situations [13]. Due to the difficulty of detecting peaks, the loss of some of them can cause the spectrum to match the background, leading to a poor estimate of the latter. Filtering based on the Fourier transform, one of the frequency filtering methods, requires separating the spectrum's frequency components from the background and noise [14].

Due to its simplicity and convenience, the polynomial approximation has become the most popular method for removing fluorescence background in a wide variety of situations. However, a manual polynomial approximation requires user intervention to select areas where the curves coincide with the original data (for determining the upper and lower limits of the frequency range), not only does this take a lot of time, but the limits change from case to case [15,16].

And although automatic polynomial approximation eliminates human intervention, their use is limited due to the high noise level. Wavelet transform methods can also be used with automated curve fitting. Still, due to the difficulty of choosing an appropriate wavelet transform threshold and the required level of resolution, the baseline display can affect the background removal results [17].

To obtain a "pure" spectrum, it is necessary to subtract the original filtered and smoothed signal's autofluorescence background. The background can be modelled by a polynomial (Equation 4), the order of which is selected based on effectively removing it and at the same time minimizing the removal of peaks in the spectrum signal [7].

\[ p(x) = p_1 x^n + p_2 x^{n-1} + \cdots + p_n x + p_{n+1}. \]  

(4)

The modified polynomial algorithm is based on successive approximation, so that the first polynomial differs from polynomial \((i - 1)\) with less than some value of \(\varepsilon\). It is generally accepted that spectra are identical if they differ by no more than 5% \((\varepsilon < 0.05)\) (Equation 5):

\[ \sigma = |\sigma_{i-1} - \sigma_i| / \sigma_i < \varepsilon, \]  

(5)

where \(\sigma_{i-1}\) is the standard deviation of the \((i - 1)\) polynomial; \(\sigma_i\) is the standard deviation of the \(i\) polynomial, now in more detail on this method. It all starts with an approximation with a single polynomial \(p_1(v)\) according to formula (Equation 4), using the "raw" signal of light \(O_0(v)\), where is the displacement of the spectrum, then the standard deviation \(\sigma\) is calculated by the formula (Equation 6),

\[ \sigma = \sqrt{(R(v_1) - R)^2 + (R(v_2) - R)^2 + \cdots + (R(v_2) - R)^2 / n} \]  

(6)

Removal of peaks is organized according to the following principle. To minimize the distortion of the approximating polynomial, the main peaks are determined based on the inequality the original signal is greater than the polynomial and the standard deviation. Now the "clean" spectrum is obtained using the difference between the polynomial approximation from the original signal [8] (Equation 7).

Where \(R(v) = O(v) - P(v)\)

(7)

Is the difference between the approximation of the polynomial and the original signal, and \(\bar{R}\) is the average value of these differences. And after that, and after that, the principal peaks are determined based on the inequality \(O(v) > p_1(v) + \sigma_1\). Points corresponding to the major peaks are removed from the spectrum and are not taken into account in subsequent steps. Peak removal is necessary to prevent unnecessary outliers in the data. The importance of this study is due to the need to study the spectral properties of aromatic compounds. The work course's expected practical results are to develop an algorithm for processing spectral data with a low signal-to-noise ratio. Also, the method of modelling the spectra of aromatic compounds and the comparative analysis of methods and the developed algorithm will help researchers complete their work.
3. Results

3.1. Spectral Characteristics of Aromatic Compounds

U.V. absorption spectrometry is a well-known analytical technique and has been applied widely in analytical chemistry [18]. Absorption is an attenuation of electromagnetic radiations at specific energy as it passes using an analyte. Electromagnetic the radiation with high energy (lower wavelength) will cause an electron by lower energy state to be able to high energy state. Beer-Lambert law is a constrained law valid for minimal concentrations of an analyte and for a monochromatic beam, provided that typically the physical and chemical attributes of the absorbed varieties tend not to change with the concentration. In absorption spectroscopy, the long optical route length and strong intake cross-sections are usually associated with high sensitivity. Nevertheless, sensor saturation may occur, which often may limit sensitivity and even linearity. Marcus et al. [19].

3.2. Analyzing Initial Data

The spectra of the sensor signal obtained with the data logger are a set of readings - pairs of numbers corresponding to the wavelength between time and the intensity of the background radiation from the sample. Time data in such a range are recorded in autofluorescence and in the phenomena that precede Figure 3. For further data, processing, we select the range of interest, according to which, after removing autofluorescence, we will select distinct peaks. In this study will be dedicated only to Benzene, Toluene, and Xylene. The spectrum of aromatic compounds understudy in the range of interest is shown in Figure 4.

Figure 3. General view of all samples of output data from the device

Figure 4. The spectrum of aromatic compounds in the range of interest obtained at a wavelength (460 nm)
3.3. Analyzing the Spectral Properties of Aromatic Compounds in Gases Using the Improved Method

As mentioned previously, the nature of noise filtering in signals is varied, so we will, can use one of the two filters, remove the noise (mainly the pulse) using an [intermediate filter] or noise filtering the spectrum using the [Savitsky-Golay] filter, shown in Figure 5.

![Figure 5](image1.png)

**Figure 5.** The spectrum of aromatic compounds in the range of interest obtained at a wavelength (460 nm) (top) before smoothing (bottom) after smoothing.

The signal from the output of the spectrograph \( O(\lambda_n) \) is represented by some additive model, consisting of the scattering signal \( RS(\lambda_n) \), the background of autofluorescence \( P(\lambda_n) \) (usually a smoothly changing line) dotted line, shown in Figure 6.

![Figure 6](image2.png)

**Figure 6.** Signal without noise (smoothed line) with a highlighted autofluorescence background \( p(\lambda_n) \) (dashed line)

To obtain a "pure" spectrum, it is necessary to subtract the original filtered and smoothed signal's autofluorescence background. The background can be modelled by a polynomial (Equation 4), the order of which is selected based on effectively removing it and at the same time minimizing the removal of signal peaks [10]. Based on our experiments, 7th and 9th order polynomials give the best approximation, shown in Figure 7.
To apply this proposed and developed method (Processors Gases), we studied and applied it to the components of aromatic compounds separately.

This is an iterative algorithm based on sequential approximation until the \((i-th)\) polynomial differs from the polynomial \((i-1)\) by less than some value \(\varepsilon\). Suppose there are points in the current spectrum that exceed the corresponding polynomial approximation considering the standard deviation. In that case, they are replaced by the polynomial points with the mean square error, thereby taking into account noise effects and avoiding artificial peaks; otherwise, the spectrum points remain unchanged \[15\]. This method considers the distortion of the signal noise and the effect of the peaks of clean, large aromatic compounds on the polynomial approximation. Still, the first and second iterations add extra parts - from removing the peaks and so on, to prevent unnecessary outliers in the data \[7\].

At the second and subsequent iterations, we perform the same procedure until the condition \((7th)\) is satisfied \[20\]. The polynomial obtained at the last iteration will be considered the autofluorescence background. A clean spectrum is obtained by subtracting this polynomial approximation from the original signal \[10\].
In Figure 8. It was observed that there was a very large discrepancy between the benzene related signal in (upper part) that was pre-treated and the pure signal obtained through the use of the algorithm in (lower part), noting that the autofluorescence signal was identified and removed afterwards, to obtain Pure signal.

**Figure 9.** Smoothed spectra with a highlighted autofluorescence background (top) and "pure" spectra (bottom) for polynomial (7th) from the signal spectrum aromatic compounds for toluene

In Figure 9. A slight discrepancy was observed between Tulane's signal in (upper part) that was previously processed and the pure signal obtained through the algorithm's use in (lower part), noting that autofluorescence was identified and then removed to obtain a pure signal. The indication in Figure 10. of pure xylene was remarkably distinct.

**Figure 10.** Smoothed spectra with a highlighted autofluorescence background (top) and "pure" spectra (bottom) for polynomial (7th) from the signal spectrum aromatic compounds for xylene
Figure 11. The scattering of the deviations from the reference value (before processing and using an average filter) for each portion of aromatic compounds (benzene, toluene, xylene) independently

Figure 12. Scatter of deviations from the reference value (median filter) for each portion of aromatic compounds (benzene, toluene, xylene) independently

Figure 13. Scatter of deviations from the reference value (PolyFit) for each portion of aromatic compounds (benzene, toluene, xylene) independently
Application of Mathematical Models and Digital Filters and their Processors of Spectral Analysis for Aromatic Compounds Gas in a Fluorescent Chemical

4. Discussion and Conclusion

Our study took care of different spectra processing methods on the principal components (Aromatic compounds). The spectral properties of the compounds and the characteristic peaks were taken and examined.

This work aims to identify the best method for preserving the fundamental components from scattering spectroscopy data. In this work, various methods of removing various distortions were investigated according to the initial model of signal observation (registration), such as registration noise and background autofluorescence. The main task was to maintain an optimal balance: to minimize the loss of the necessary information and at the same time to remove all unnecessary noise.

Files with spectroscopy data from the scattering sensor are entered into the program. The raw spectra are processed (procedures Med Filter, SGF, PolyFit, ProcessorsGases), which is "pure" spectra.

Further, the obtained sets are compared with sets from a priori data. According to the comparison data, the influence of distortions introduced by various algorithms for noise reduction and background removal of autofluorescence is determined.

The scatter diagrams (whisker-plot) were constructed to introduce the principal components' introduced absolute distortions during preprocessing (stages of noise reduction and removal of autofluorescence). The Savitsky-Golay filter has a smaller error spread than the median filter. Therefore it is easier to correct this spread. And among the methods of removing autofluorescence, the improved method (ProcessorsGases) have the smallest error spread compared to other methods.

Suppose there are points in the current spectrum that exceed the corresponding polynomial approximation considering the standard deviation. In that case, they are replaced by the polynomial points with the root mean square error, thereby taking into account noise effects and avoiding artificial peaks; otherwise, the spectrum points remain unchanged. At the second and subsequent iterations, we perform the same procedure until the condition (Equation 5) is satisfied. The polynomial obtained at the last iteration will be considered the autofluorescence background. A "clean" spectrum is obtained by subtracting this polynomial approximation from the original signal.

Conflicts of interest: The authors declared no conflict of interest.

Acknowledgements: Sponsor acknowledgements Samara State University.

References


