6-2-1995

Nondispersive triple-correlation spectrometer for trace gas analysis

Samuel Peter Kozaitis

Bruce J. Bradshaw

PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Nondispersive triple-correlation spectrometer for trace gas analysis

Samuel Peter Kozaitis
Bruce James Bradshaw

SPIE.

Nondispersive triple correlation spectrometer for trace gas detection

S. P. Kozaitis and B. J. Bradshaw
Florida Institute of Technology
Division of Electrical, Computer Science, and Engineering
150 W. University Blvd.
Melbourne, FL 32901-6988

ABSTRACT

We showed that the output of an autocorrelation with a gas spectrum could be triple-correlated to reduce the effects of Gaussian noise. Such processing is directly applicable to a correlation spectrometer but could be applied to other designs. In spite of the added processing, triple-correlating the output of a spectrometer allowed us to generally increase the SNR by over a factor of ten. Even with as few as two measurements good results were obtained.

Keywords: correlation, higher-order statistics, triple-correlation, spectrometer

1.0 INTRODUCTION

The need to detect atmospheric gases is often related to concerns about the environment. Also, since a large portion of our energy imports end up as fuels for use in the generation of steam and electricity, the monitoring of the efficiency of combustion in power generation situations is an economical, environmental, and conservation related necessity.

Gas measurements of the atmosphere can be performed by many methods but one of the best known is by infrared absorption. Using a correlation process with a matched filter of the spectrum of the gas of interest often rejects much of any nearby absorption caused by another gas species; however, the correlation process can be sensitive to noise.

In this paper we applied higher-order correlation methods to correlation-based methods of spectral detection such as with a gas-filter spectrometer.[1,2] Such higher-order correlations have been theoretically shown to eliminate Gaussian noise of unknown spectral density under certain conditions. The noise elimination is intimately related to the properties of higher than second-order moments of the Gaussian

0-8194-1831-5/95/$6.00

probability function. This property causes noise of Gaussian form in measurements to statistically approach zero.

2.0 HIGHER-ORDER STATISTICS

The correlation between two discrete functions $a(k)$ and $b(k)$ is defined as

$$c(t) = \sum_{k=0}^{N-1} a(k)\, b(t+k) \tag{1}$$

where t ranges from 0 to N - 1, and N is the number of samples contained in one period. Using Eq. (1) a comparison between two spectra can be performed. When considering absorption spectra, usually t is set to zero and a correlation is performed between an unknown spectrum and a known spectrum. The magnitude of this result when compared to the magnitude of the autocorrelation of the known spectrum can be used for classification of the unknown spectrum.

2.1 Higher-order correlations

The correlation function as described in Eq. (1) is only one of many calculations which can be used as a statistical tool. In recent years higher-order correlations have found new interest in several areas. This is due primarily because in some cases such as the Gaussian noise distribution, the noise contribution can be theoretically eliminated under certain conditions.

The nth-order correlation of the signal $a(k)$ is defined as[3]

$$a_n(t_1, t_2, \ldots t_{n-1}) \int \sum_{k=0}^{N-1} a(k)\, a(t_1+k)\, a(t_2+k) \ldots a(t_{n-1}+k) \tag{2}$$

where the nth-order correlation is a function of n - 1 independent variables. Spectral comparisons are most useful at zero displacement. From Eq. (2), the zeroth lag of the nth-order correlation can be written as

$$a_n(0, 0, \ldots 0) = \sum_{k=0}^{N-1} a^n(k) \quad . \tag{3}$$

Therefore, the *n*th-order correlation of $a(k)$ at zeroth lag is equal to the sum of the *n*th power of $a(k)$.

Because the *n*th-order correlation is a function of n - 1 variables, the n = 3, (third-order or triple correlation) of a 1-D function is a function of two variables. The triple correlation of $a(k)$ is symmetric with respect to its variables. There is no distinction between the variables, and their reversal should yield the same function. For example, Fig. 1(a) indicates a function $a(k)$, and its triple correlation is shown in Fig. 1(b). Note that if t1(t2) is set to zero the second-order correlation, or autocorrelation appears on the t2(t1) axis.

2.2 Cumulants

Higher-order processes are defined in terms of cumulants. For orders n = 1, 2, 3, the cumulants $c_n(t_1, t_2, \ldots t_n - 1)$ of the function $a(k)$ are related to correlations by,

$$c_1 = a_1 \tag{4}$$

where a_1 is the mean of $a(k)$,

$$c_2(t_1) = a_2(t_1) - (a_1)^2 \tag{5}$$

where $a_2(t)$ is the autocorrelation of $a(t)$, and

$$c_3(t_1, t_2) = a_3(t_1, t_2) - a_1[a_2(t_1) + a_2(t_2) + a_2(t_2 - t_1)] + 2(a_1)^3 . \tag{6}$$

If the function $a(k)$ is zero-mean ($a_1 = 0$), then the second and third-order correlations are identical to the second and third-order cumulants. The general relationship between correlations and cumulants can be found in Ref. 4.

Given T independent records of the sequence $a(k)$, the cumulants can be estimated by averaging as

$$\hat{c}_n(t_1, t_2, \ldots t_n\text{-}1) = \frac{1}{T} c_n(t_1, t_2, \ldots t_n\text{-}1) \quad (7)$$

However, if the records follow a Gaussian distribution, then for $n \geq 3$, $\hat{c}_n(t_1, t_2, \ldots t_{n-1})$ will approach zero if T is large enough. Therefore, if a zero-mean signal is corrupted by Gaussian noise, then by averaging T records of the noisy signal, the contribution to the cumulant due to noise will approach zero. Furthermore, if the samples follow any symmetric distribution, for $n = 3$ the cumulants will approach zero. Therefore, the cumulants will correspond to only the signal and will not have any contribution from noise.

3.0 SPECTRAL DETECTION USING CORRELATIONS

We described an analytical example of spectral comparison using both the second-and third-order correlation techniques. We considered the result of a noisy spectrum that was autocorrelated with a reference spectrum with little noise. We assumed that the noisy spectrum contained zero-mean noise of unknown spectral density. The result of the noisy spectral correlation ($n = 2$), was triple-correlated. The reference spectrum was described by $a(k)$, and the noisy input spectrum was described by $a(k) + v(k)$, where $v(k)$ indicates noise. The second-order correlation between the reference and noisy input was described as $y(k)$.

Using Eq. (3) the second-order correlation of $y(k)$ at zero displacement was written as

$$y_2(0) = \sum_{k=0}^{2N-1} y^2(k) \quad (8)$$

We averaged $y_2(0)$ over T records and wrote the result as

$$\hat{y}_2(0) = E\left\{ \sum_{k=0}^{2N-1} a^2{}_2(k) + \sum_{k=0}^{2N-1} y^2{}_v(k) + 2\sum_{k=0}^{2N-1} a_2(k) y_v(k) \right\} \tag{9}$$

where $y_v(k)$ is the result of the second-order correlation of the noise and the reference spectrum. Because we assumed zero-mean noise, the last term is in Eq. (9) vanishes. Furthermore, noise is widely assumed to be an ergodic process; therefore, Eq. (9) became

$$\hat{y}_2(0) = \sum_{k=0}^{2N-1} a^2{}_2(k) + E\left\{ y^2{}_v(k) \right\} \tag{10}$$

From Eq. (10) it is clear that as the noise power increases, the SNR decreases.

We also processed the result of the noisy spectral comparison using the triple correlation. Using Eq. (3) the triple correlation of $y(k)$ at zero displacement was written as

$$y_3(0, 0) = \sum_{k=0}^{2N-1} y^3(k) \tag{11}$$

We averaged $y_3(0,0)$ over T records and wrote the result as[5]

$$\hat{y}_3(0,0) = E\left\{ \sum_{k=0}^{2N-1} a^3{}_2(k) + \sum_{k=0}^{2N-1} y^3{}_v(k) + 3\sum_{k=0}^{2N-1} a^2{}_2(k) y_v(k) + 3\sum_{k=0}^{2N-1} a_2(k) y^2{}_v(k) \right\} \tag{12}$$

Because we have assumed zero-mean noise, the second term is in Eq. (12) is equal to the third-order cumulant of $v(i)$ and vanishes if the noise is Gaussian. The third term is multiplied by the average of $v(i)$, and for T large enough will be zero. After separating the factors in the fourth term Eq. (12) then became[5]

$$\hat{y}_3(0,0) = \sum_{k=0}^{2N-1} a^3{}_2(k) + 3E\left\{ y^2{}_v(k) \right\} \sum_{k=0}^{2N-1} a_2(k) \tag{13}$$

where the second term in Eq. (13) is a product of signal and noise.

4.0 RESULTS

We considered a binary hypothesis where we had to decide between the presence of a spectrum or noise. For the second-order correlation the result of Eq. (10) can be compared to the quantity $E\left\{ y^2{}_v(k) \right\}$. However, as the noise increases the ratio of the first term to the second term in Eq. (10), the SNR, decreased. Therefore, the difference between Eq. (10) and $E\left\{ y^2{}_v(k) \right\}$ would also decrease. In the case of the triple-correlation, the result in Eq. (13) can be compared to zero for classification. Or, if T is not large enough for the noise to be completely eliminated, Eq. (13) can be compared to $E\left\{ y^3{}_v(k) \right\}$. The second term in Eq. (13) is the product of a signal and noise term. Therefore, the difference between Eq. (13) when signal plus noise is present and when only noise is present, increases as the noise increases.

In Eq. (13), because of the noise-signal product term the SNR is ambiguous; therefore, we determined the SNR statistically using a general definition of the SNR applicable to practical measurements which was written as[6]

$$\text{SNR} = \frac{\left(\text{E}\{\hat{y}(0)\}\right)^2}{\text{var}\{\hat{y}(0)\}} \tag{14}$$

where E{} and var{} are the expected value and variance respectively, and $\hat{y}(0)$ is the measured value of interest.

We considered the detection of benzene whose spectrum[7] is shown in Fig. 2. We performed a correlation between a noiseless spectrum and one with noise added using 256 length sequences. Then, the triple-correlation of the result was calculated for different amounts of input noise and different numbers of iterations. In addition, we repeated every experiment 10 times and averaged the results. From this data we calculated the output SNR in Eq. (14). The results corresponding to an input SNR of 0.01 - 10 are shown in Fig. 3.

The results showed that the SNR increased dramatically when the result of the noisy correlation was triple-correlated. In addition, the increase in SNR was clear even when using a small number of iterations. Furthermore, we used only noise as an input and recorded the variance of the triple-correlation. These results are shown in Table 1. The results show that the variance is generally small when just noise is present. This indicates that a threshold may be chosen near zero to confidently detect benzene.

5.0 CONCLUSION

We showed that the output of an autocorrelation with a gas spectrum could be triple-correlated to reduce the effects of Gaussian noise. Such processing is directly applicable to a correlation spectrometer but could be applied to other designs. In spite of the added processing, triple-correlating the output of a spectrometer allowed us to generally increase the SNR by over a factor of ten. Even with as few as two measurements good results were obtained.

6.0 REFERENCES

[1] R. M. Goody, "Cross-Correlating Spectrometer," *J.Opt Soc. Am.*, 58, 900 (1968)

[2] H. H. Zwick, R H. Weins, *Infrared, Correlation, and Fourier Transform Spectroscopy,* J. Mattson, H. B. Mark Jr., H. C. Mac Donald Jr. Eds., Marcel Dekker Inc. (1977)

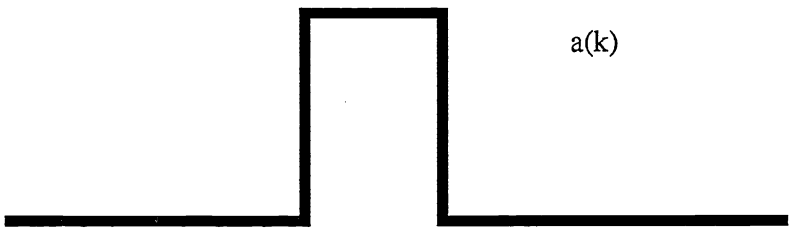
[3] J. G. Proakis, C. M. Rader, F. Ling, and C. L. Nikias, *Advanced Digital Signal Processing*, Macmillan: New York (1992)

[4] M. Rosenblatt, *Stationary Sequences and Random Fields*, Birkhauser: Boston, MA (1985)

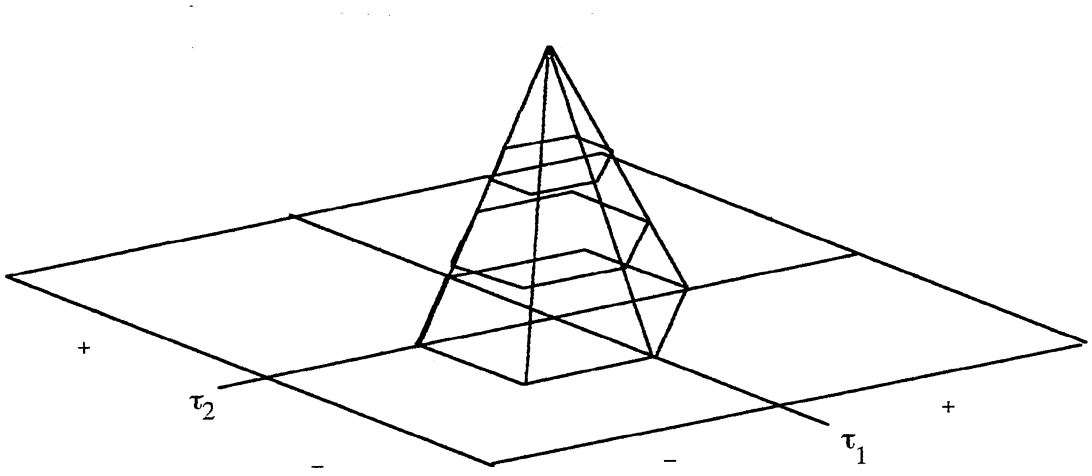
[5] G. B. Giannakis, and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol 38, no. 7, 1284-1296 (1990)

[6] A. J. Viterbi, *Principles of coherent communication*, McGraw-Hill: New York (1966)

[7] R. H. Pierson, A. N. Fletcher, and E. S. Gantz, "Catalog of infrared spectra for qualitative analysis of gases," *Anal. Chem*, vol. 28, no. 8, 1218-1239 (1956)



(a)



(b)

Figure 1 Triple correlation example (a) an example function (b) its triple correlation

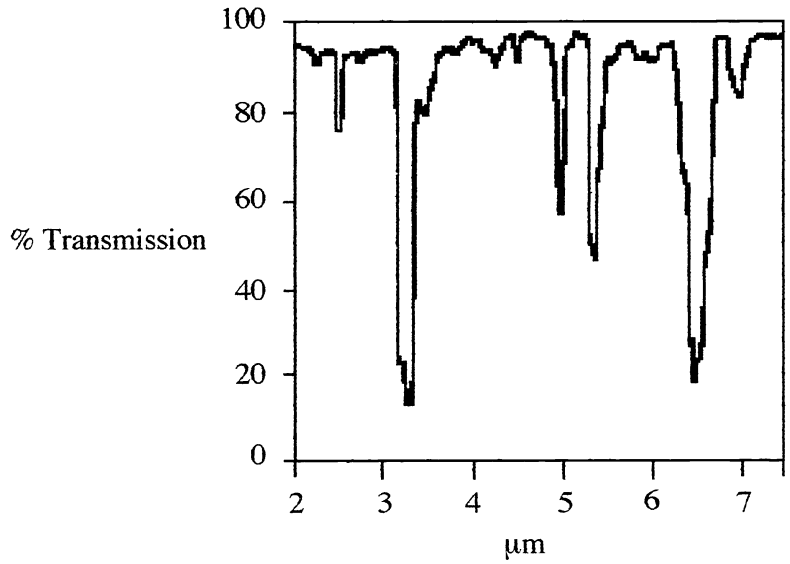


Figure 2 Spectrum of Benzene

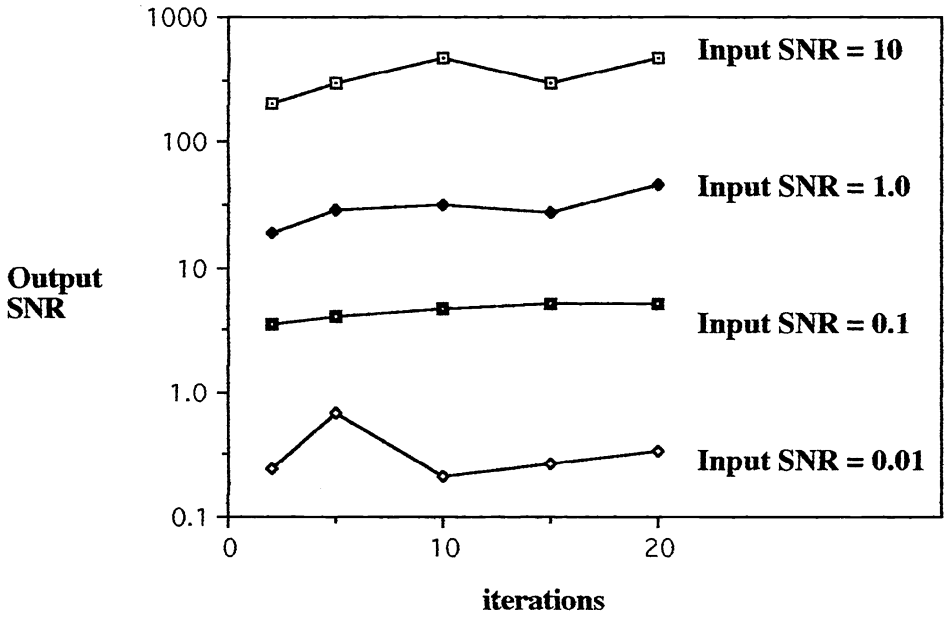


Figure 3 Graph of output SNR vs. number of iterations of different input SNRs

Table 1 Output variance when input was Gaussian noise

Input SNR	iterations				
	2	5	10	15	20
0.01	0.78	0.47	1.06	0.54	.75
0.1	0.68 E-3	0.31 E-3	0.40 E-3	0.67 E-3	0.31 E-3
1.0	0.80 E-6	0.27 E-6	0.54 E-6	0.22 E-6	0.21 E-6